普通高校物联网工程专业规划教材

物联网安全

任 伟 编著

清华大学出版社

物联网安全

任　伟　编著

清 华 大 学 出 版 社
北　京

内 容 简 介

本书全面而又系统地论述了物联网安全中的部分关键问题及其典型解决方案。全书分为 3 大部分：物联网感知层安全、物联网网络层安全和物联网应用层安全。物联网感知层安全介绍 RFID 安全、无线传感器网络安全、物联网终端系统安全；物联网网络层安全介绍近距离无线接入安全（无线局域网安全）、远距离无线接入安全（无线移动通信安全）、接入网安全的扩展讨论、物联网核心网安全（6LoWPAN 安全和 RPL 安全）、物联网服务端安全（云计算安全）；物联网应用层安全介绍智能电网安全、EPCglobal 网络安全、基于无线体域网的远程医疗安全、M2M 安全。

本书可作为物联网工程、信息安全、计算机科学等专业的研究生或本科高年级教材，对物联网安全领域的研究者具有一定参考价值，对物联网领域的工程技术人员亦具有指导价值。

图书在版编目（CIP）数据

物联网安全/任伟编著. —北京：清华大学出版社，2012. 6
（普通高校物联网工程专业规划教材）
ISBN 978-7-302-28503-8

Ⅰ. ①物…　Ⅱ. ①任…　Ⅲ. ①互联网络－应用－物流 ②互联网络－安全技术　Ⅳ. ①TP393. 4 ②F253. 9

中国版本图书馆 CIP 数据核字(2012)第 064979 号

责任编辑：龙启铭
封面设计：傅瑞学
责任校对：李建庄
责任印制：王静怡

出版发行：清华大学出版社
网　　址：http://www. tup. com. cn，http://www. wqbook. com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup. tsinghua. edu. cn
质量反馈：010-62772015，zhiliang@tup. tsinghua. edu. cn
课件下载：http://www. tup. com. cn，010-62795954
印 刷 者：北京市世界知识印刷厂
装 订 者：三河市溧源装订厂
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：12. 75　　**字　　数**：315 千字
版　　次：2012 年 6 月第 1 版　　**印　　次**：2012 年 6 月第1 次印刷
印　　数：1～3000
定　　价：25. 00 元

产品编号：046006-01

前　言

物联网作为国家的战略新兴产业，正在得到大力发展，在国家科技创新、可持续发展和产业升级中具有重要的地位。在工业和信息化部最近发布的《物联网"十二五"发展规划》中，"加强信息安全保障"已成为主要任务之一。但是，目前还没有全面论述物联网安全的相关书籍，因此，无论是信息网络安全领域的科研与教学人员，还是物联网工程领域的技术人员，都急需一本全面讲解物联网安全的书籍。

物联网安全本身涉及的内容极其广泛，本书特精心挑选了其中的关键问题、特色问题、重点问题进行讨论，给出相关的技术、方法和常用的典型方案，并介绍几个特色领域（如 6LoWPAN、智能电网、EPCglobal 网络以及 M2M 领域）的安全研究进展。

作者在写作的过程中特别遵循了以下一些新的思路。

（1）内容编排循序渐进、由浅入深、先总体再局部、兼顾广度和深度。先给出物联网的体系结构以及安全架构，再给出物联网安全问题的共性之处和一般解决思路。然后根据物联网的安全架构（感知层安全、网络层安全和应用层安全），分章节依次探讨特定的安全问题。同时在每一章的介绍中，还遵循先介绍总体架构与安全威胁全貌，后论述典型安全论题的次序。

（2）选材新颖、理论联系实际。选材尽量突出基本的研究问题以及新的进展，理论的论述突出共性和一般原理（如对接入网围绕认证与密钥协商机制来论述、RFID 认证协议、无线传感器网络的密钥管理），实践部分强调新颖性和工程性（如对最新 LTE 加密算法 ZUC、LBlock、NTRU、SMS4、SM3 的解释）。理论与实际结合中，突出网络安全设计方案的一般规律，便于指导将来的安全设计。

（3）注重启发性和对创新能力的培养，包括对一般原理的总结和归纳、协议设计方法的比较和分析，注重对问题本质的提炼。通过每一章的"研究与思考"环节，启发读者思考、提出并解决新的物联网安全问题。爱因斯坦曾经说过："提出问题往往比解决问题更为重要"。这一环节对于创新教育可能是一次有益的尝试。

（4）注重对国内自主知识产权和自主创新的介绍，包括轻量级分组加密 LBlock、国家密码管理局标准 Hash 算法 SM3，无线局域网安全标准 WAPI、流密码 SMS4、3G 标准 TD-SCDMA、4G 中的流加密算法 ZUC、国产手机操作系统 OMS、北斗卫星导航系统、CMMB 等。由于信息安全的行业特殊性和我国综合国力的提升，介绍这部分相关成果有利于激发读者和相关技术人员对我国自主创新成果的关注，扩大我国自主知识产权成果的影响，促进我国自主知识产权成果的推广和应用。

（5）特别讨论了物联网安全中的几个特色问题，如 6LoWPAN 的安全、智能电网安全、EPCglobal 网络以及 M2M 安全。对从事该方面研究的科研人员以及安全工程技术研发人员有一定的参考价值。

(6) 注重对实践能力的培养和对行业动态的关注。书中给出了多个密码算法便于读者去实现。通过文中对最新物联网科技动态的报道(如 iPhone 集成 RFID 的专利信息),便于把握新的研究动向和应用场景。文中还对国内部分行业领先企业做了介绍,以支持民族产业和扩大自主品牌的影响。

(7) 密码学工具与计算机和网络安全的有机结合。物联网安全问题的解决工具主要是密码学,因为它深刻而且精巧,但是又不囿于密码学方法(有些问题需要来自计算机科学的方法解决)。实际中的网络问题为密码学提供了丰富的问题和需求,刺激了密码学的发展。同时,根据不同的安全问题,应选择与之相适应的安全方案,而不能局限于某一种方法或者工具。本书在两种方法之间力求兼顾,希望能给有密码学背景的读者提供一些实际问题,也给计算机科学背景的读者提供一些密码学工具和密码学思考问题的角度(特别是在云存储安全一节)。

全书共分 13 章:第 1 章是物联网安全概述,第 2 章介绍 RFID 安全,第 3 章介绍无线传感器网络安全,第 4 章介绍物联网终端系统安全,第 5 章介绍近距离无线接入安全——无线局域网安全,第 6 章介绍远距离无线接入安全——无线移动通信安全,第 7 章介绍接入网安全的扩展讨论,第 8 章介绍物联网核心网安全——6LoWPAN 和 RPL 的安全性,第 9 章物联网服务端安全——云计算安全,第 10 章介绍智能电网安全,第 11 章介绍 EPCglobal 网络安全,第 12 章介绍基于无线体域网的远程医疗安全,第 13 章介绍 M2M 安全。其中,第 2~4 章为第 1 部分物联网感知层安全,第 5~9 章为第 2 部分物联网网络层安全,第 10~13 章为第 3 部分物联网应用层安全。带星号的部分为选学内容。

本书面向的主要对象包括从事信息和网络安全研究的科研人员,学习物联网安全相关课程的高等院校信息安全类、物联网工程类、计算机类专业研究生和本科高年级学生,以及从事物联网安全技术研发、应用和管理的工程技术人员。

本书得到了国家自然科学基金面上项目(No. 61170217)、湖北省高等学校省级教学研究项目(2011123,2011125)、山东省计算机网络重点实验室开放课题资助(SDKLCN-2011-01)、中央高校基本科研业务费专项资助项目(090109,110109)的支持,在此表示感谢。感谢长江学者华中科技大学金海教授和长江学者北京邮电大学杨义先教授的指导和教诲。感谢新加坡管理大学邓慧杰(Robert H. Deng)教授的帮助。感谢香港科技大学和美国伊利诺理工大学的老师和朋友们。感谢研究生刘宇靓协助绘制了部分插图。

愿本书的写作能为我国物联网安全的研究以及教学起到抛砖引玉的作用。由于作者水平和学识有限,加之编写时间紧,不足之处在所难免,在此衷心恳请广大读者同行批评指正。

任　伟

中国地质大学(武汉)

2012 年 4 月

目　　录

第 3 部分　物联网应用层安全

第 1 章　物联网安全概述

1.1　物联网安全概述

1.1.1　物联网概念与发展历程

1. 物联网的概念

物联网(Internet of Things)是一个基于互联网、传统电信网等信息承载体,让所有能够被独立寻址的普通物理对象实现互联互通的网络[1]。它具有普通对象设备化、自治终端互联化和普适服务智能化 3 个重要特征。

另外一个在国内被普遍引用的物联网定义是来自百度百科和互动百科的定义,虽然没有经过官方审定,但是传播范围很广:通过 RFID、红外感应器、全球定位系统、激光扫描器等信息传感设备,按约定的协议,把任何物品与互联网连接起来,进行信息交换和通信,以实现智能化识别、定位、跟踪、监控和管理的一种网络。

下面先回顾一下物联网的发展历程,然后介绍如何理解物联网概念。括号部分是我们对事件的评论。后面的章节选材和内容组织在某种程度上与这些事件相呼应。

(1) 1998 年,美国麻省理工学院提出了当时被称为 EPC(Electronic Product Code)系统的物联网构想。紧接着在 1999 年,在 EPC 编码、RFID 技术和互联网基础上,MIT 的 Auto-ID 中心提出物联网的概念。2003 年 10 月,非盈利性组织 EPCglobal 成立(这形成了基于 Internet 的 RFID 系统)。

(2) 2004 年,IETF 成立了基于低功耗无线个域网(LoWPAN)的 IPv6 工作组 6LoWPAN,致力于研究在由 IEEE 802.15.4 链路构成的低功耗无线个域网中如何优化运行 IPv6 协议。这为通过 Internet 直接寻址访问无线传感器网络结点(无需通过网关)提供了可能(使得无线传感器网络走向开放并可能成为一种 Web 服务)。

(3) 2006 年,美国国家自然科学基金委员会将信息物理融合系统 CPS(Cyber Physical System)作为重点支持的研究课题。CPS 是一个以通信和计算为核心的集成的监控和协调行动的工程化物理系统,是计算、通信和控制的融合,具备很高的可靠性、安全性和执行效率。CPS 试图突破原有传感器网络系统自成一体、计算设备单一、缺乏开放性等缺点,注重多个系统间的互联互通,强调与互联网的联通,真正实现开放的、动态的、可控的、闭环的计算和服务支持(感知和控制融合使得物联网更加强大,从此控制系统的安全需要重视了)。

(4) 2005 年,国际电信联盟 ITU 发布了《ITU 互联网报告 2005:物联网》。报告指出,世界上所有的物体都可以通过互联网主动进行信息交换。RFID、传感器技术、纳米技术、智能嵌入技术将得到更加广泛的应用,强调 M2M(Machine-to-Machine)通信。2008 年,欧洲智

能系统集成技术平台(EPoSS)在《物联网 2020》报告中分析预测了未来物联网的发展阶段(可见,欧洲的物联网是从电信部门开始主导的,因为 M2M 具有巨大的市场潜力)。

(5) 2008 年 9 月,IPSO(IP Smart Object)联盟成立,推进 IP 在智能物体(Smart Object)中的应用(智能物体可视为一种通用的物联网终端模型,其功能可能是异构的 Hybrid,即具有感知、识别、制动等多重功能)。

(6) 2009 年 1 月,奥巴马就任美国总统后,与美国工商业领袖举行了一次"圆桌会议"。作为仅有的两名代表之一的 IBM 首席执行官首次提出了"智慧地球"的概念。依据奥巴马总统的经济恢复法案,2009 年美国能源部宣布投资 45 亿美元打造基于 M2M 技术的实时双向通信的智能电网。在美国除 M2M(Machine-to-Machine)外,最受关注的物联网应用是智能电网和远程医疗。这两个领域都是奥巴马政府低碳经济和医疗改革政策直接推动的结果(美国研究物联网是从具体应用入手的,重视智能电网、远程医疗等物联网应用)。

(7) 2009 年,欧盟执委会发表题为《Internet of Things-an Action Plan for Europe》的物联网行动方案,描绘了物联网技术应用的前景(物联网上升为整个欧盟的战略行为)。

(8) 2009 年,韩国通信委员会和日本政府 IT 战略本部分别提出了物联网相关战略(韩日的物联网战略)。

(9) 2009 年 8 月,温家宝总理在无锡视察时发表重要讲话,提出"感知中国"的战略构想。在后来的"让科技引领中国可持续发展"的讲话中,将物联网列入战略新兴产业之一,标志着物联网产业发展已经提升到我国的国家战略(我国开始大规模介入物联网)。

对于物联网概念的理解,应该从基本应用需求出发,把握物联网的特点和基本技术。对于物联网的发展和应用,不同的国家有着不同的着力点。美国常常提及智能电网(为了新能源利用、节能减排的需要)、远程医疗、基于 EPCglobal 网络的供应链管理。欧洲常常提及 M2M 应用,从大规模安装无线移动通信的 SIM 卡到智能设备或者监控仪表来促进其发展,以及智能嵌入式实时系统。我国的物联网则涉及范围更加广泛,从传感网和 RFID 的应用入手,到两化融合(自动化和信息化的融合)和 M2M,被认为是一次赶超世界先进信息技术的历史机遇。

物联网的特点是融合了无线网络和有线网络,扩大了接入 Internet 网络的设备的规模(除了计算机外,还有大量的微型计算设备),使得网络连接的范围更广。加上这些设备如传感器结点具有感知外部环境的功能,有些设备如 RFID 具有标识附着的物体的能力,这些设备还可以借助卫星定位系统如 GPS,可被定位和追踪,这些都使得人类具有比以前更加强大的获取信息的能力。如果这些设备还能够具备行动能力,则人类具有比以前更加强大的控制能力。这些使得人类具有前所未有的能力去感知、标识、跟踪、联接、控制、管理地球上的物体的一举一动,好比给地球加上了一个神经系统。这个神经系统有末梢(物联网终端系统)、有传导(网络通信系统)以及处理(如云计算、信息与网络中心等)。

1.1.2 物联网的体系结构

物联网应该具备 3 个特征:

(1) 全面感知,即利用 RFID、传感器、条形码(二维码)、GPS(北斗卫星导航)定位装置等随时随地获取物体的信息。

(2) 可靠传递,通过各种网络与互联网的融合,将物体的信息实时准确地传递出去。

(3) 智能处理,利用云计算等各种智能计算技术,对海量数据和信息进行分析和处理,对物体实施智能化控制。因此,物联网的体系架构通常认为有 3 个层次:底层是用来感知(识别、定位)的感知层,中间是数据传输的网络层,上面是应用层,如图 1.1 所示。

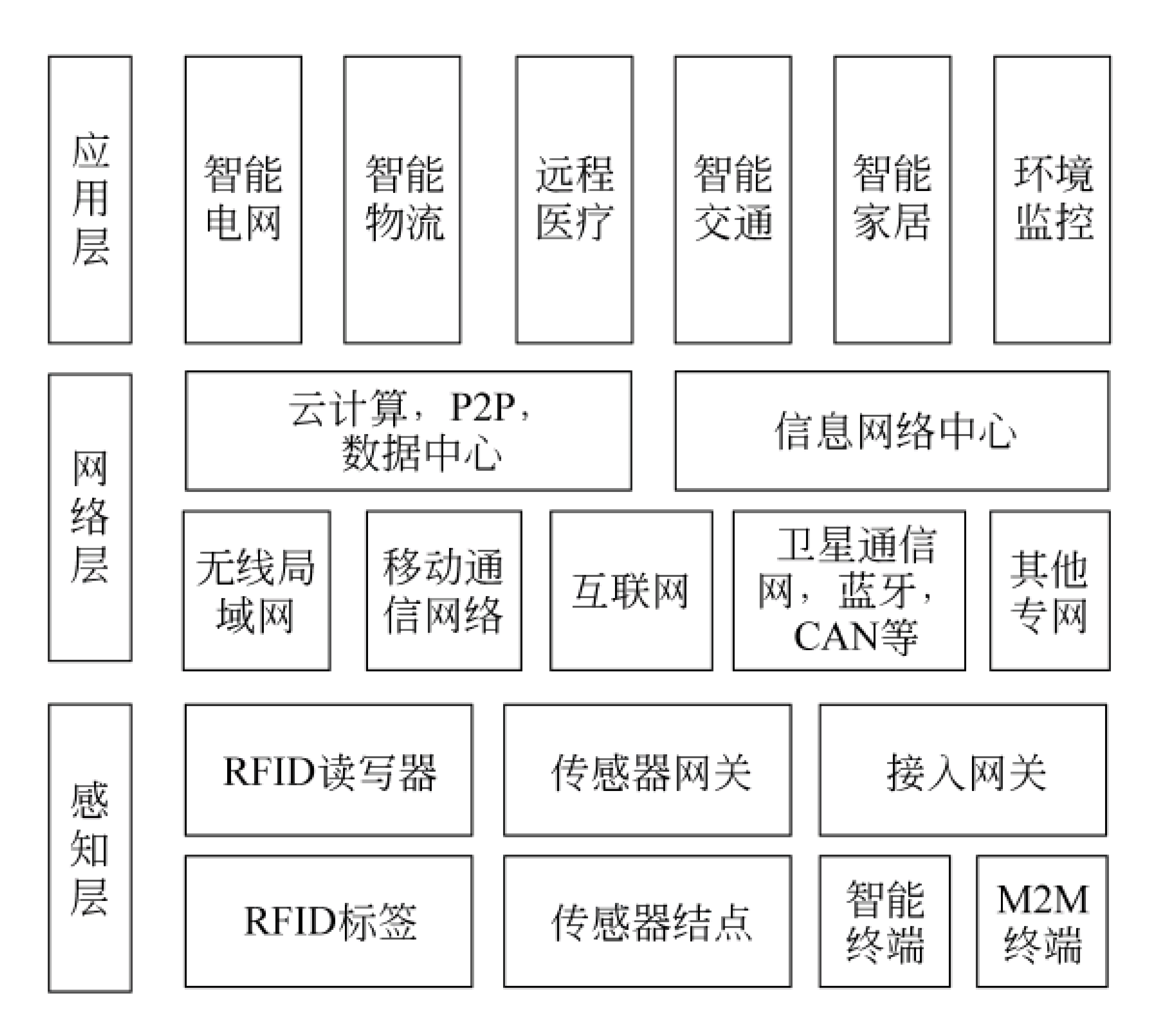


图 1.1　物联网体系架构

感知层包括以传感器为代表的感知设备、以 RFID 为代表的识别设备、GPS(北斗系统)等定位追踪设备以及可能融合部分或全部上述功能的智能终端(手机)等。大规模的感知则构成了无线传感器网络。另外,M2M 的终端设备,智能物体都可视为感知层中的物体。感知层是物联网信息和数据的来源。

网络层包括接入网、核心网以及服务端系统(云计算平台、信息网络中心、数据中心等)。接入网可以是无线近距离接入,如无线局域网、ZigBee、Bluetooth、红外,也可以是无线远距离接入,如移动通信网络、WiMAX 等,还可能其他接入形式如有线网络接入(PSTN、ADSL、宽带)、有线电视、现场总线、卫星通信等。网络层的承载是核心网,通常是 IPv6(IPv4)网络。网络层是物联网信息和数据的传输层,此外,网络层也包括信息存储查询、网络管理等功能。云计算平台作为海量感知数据的存储、分析平台,是物联网网络层的重要组成部分,也是应用层众多应用的基础。

应用层利用经过分析处理的感知数据为用户提供丰富的特定服务,这些服务通常是在具备感知、识别、定位追踪能力后新增加的功能,如智能电网、智能物流、远程医疗、智能交通、智能家居、环境监控等。依靠感知层提供的数据和网络层的传输,进行相应的处理后,可能再次通过网络层反馈给感知层。应用层是物联网信息和数据的融合处理和利用,是物联网发展的目的。

我们认为,物联网中比较有特色的共性网络技术有 3 个:6LoWPAN、EPCglobal 网络和 M2M(Machine-to-Machine)。

(1) 6LoWPAN,主要用于基于 Internet 寻址访问传感器结点,由 IETF 定义,被 IPSO 联盟推广。从广义上讲,可用于在基于 IEEE 802.15.4 的无线个域网链路条件下,

承载 IPv6 协议构成一个广域的大规模的设备(智能物体)的联网。这一技术可视为无线传感器网络的 Internet 演进,其推动者是 IETF 以及 IPSO 联盟。

(2) EPCglobal 网络,主要用于基于 Internet 的 RFID 系统,由 EPCglobal 定义,主要用于广域物体的定位与追踪的物流应用。这一技术可视为 RFID 技术的 Internet 演进,其推动者是 EPCglobal 组织。

(3) M2M,通常是指通过远距离无线移动通信网络(例如 GPRS、TD-SCDMA 等)的设备间的通信,如终端设备与中心服务器间通信的智能抄表,以及两个广域网的设备间的通信(通过中心服务器转接)。M2M 的主要作用是为远端设备提供无线通信接入 Internet 的能力。M2M 很多时候可视为一种接入方式,这种接入方式和无线移动通信网中以人为中心的接入方式不同,M2M 中接入的对象是设备,且这些设备通常是无人看守的(因此 M2M 设备可能是机卡一体的)。当然,广义上 M2M 可泛指所有机器之间的通信,涵盖控制系统间的通信。M2M 通常是移动通信运营商在推动,可视为远距离无线移动通信网络的接入端从以人为中心向以设备为中心演进。

上述 3 种技术之间的关系可以表述如下:

(1) EPCglobal 网络和 M2M 可以融合,即 RFID 读写器通过 M2M 连接到 Internet,然后可访问 EPGglobal 定义的 ONS(Object Name Service)、EPCIS(EPC Information Service)等服务。EPCglobal 网络主要定义了应用层服务的架构。

(2) 6LoWPAN 和 M2M 之间的区别是前者提供了直接的 Internet 寻址能力,而后者可以通过在 M2M 服务器端的网关功能进行寻址,这种寻址类似于一种基于广域无线通信网的网络地址转换(Natural Address Translation,NAT),因为后者可不需要配置 IP 地址,而只需要配置 M2M 标识。6LoWPAN 是协议栈的一个适配层。

(3) 无线传感器结点或者无线传感网网络网关也可以通过 M2M 的 GPRS、特别是 TD-SCDMA 连接到远距离无线移动通信网络的中心结点,然后与 Internet 相连。达到 6LoWPAN 技术类似的效果。M2M 的最大优势是对大规模移动性的支持。

基于上述基本网络技术,根据需求选择适当的终端设备,再合理地选择接入网络和核心网,就可以构造各种新颖的应用。

国际电信联盟(ITU)在 2005 年的物联网报告中,描述了一个物联网应用场景。这是 2020 年日常生活的一天。一个来自西班牙的 23 岁名叫 Rosa 的学生,刚刚同男朋友吵完架,想要独处一段时间。她决定私自驾驶自己的智能汽车去法国阿尔卑斯山的一个滑雪胜地度周末。但是她必须先去修理厂,因为汽车的传感系统提醒她轮胎可能坏了。当她进入修车厂入门通道的时候,基于传感器的诊断工具已经为她的车做了全面检查,并根据检查的结果引导她的车开进一个配备有自动机器人手的专门修理站点。Rosa 下车后去喝杯咖啡,饮料自动售货机知道 Rosa 喜欢冰咖啡,所以在 Rosa 挥舞网络手表付账后得到了一杯她想要的冰咖啡。当 Rosa 回来时,一对新的后轮胎(装有传感器和 RFID)已经安装好了。机器人然后提示 Rosa 新轮胎上与隐私有关的选项,存储在汽车控制系统中的信息是为维护和维修用的,但在汽车行驶中如果周围有 RFID 读写器,这些信息将被读取到。Rosa 不想任何人知道(特别是她男友)知道她去哪里,所以她选择把这些信息设为被保护,防止被无权限的人看到。

终于,Rosa 可以开车去最近的商业街购物了。她想买有内嵌媒体播放器和温度调节功能的单板滑雪服。由于她要去的滑雪场已经通过无线传感器网络监测到雪崩的可能性很小,所以她感到去那里很安全。在通过法国和西班牙边境时,她不需要停留,因为她的车里保存放了她的驾照和护照,在越过边境的时候,这些信息被自动传输到边境控制装置中自动检查放行。

突然,Rosa 从她的太阳镜上收到一个视频寻呼,她赶紧把车停到路边,她看到男友正请求原谅并问她是否想一起度周末。这时她心情正好不错,于是她脱口而出,对导航系统发出了撤销隐私保护的语音命令,这样她的男友就可以看到她现在位置赶过来。

1.1.3　物联网的安全架构

物联网的安全架构可以根据物联网的架构分为感知层安全、网络层安全、应用层安全。应用层安全的研究内容,可能会与感知层安全和网络层安全有交叉,但其关注重点是具有应用特色的安全问题,或者需要在应用层解决的安全问题,如密钥管理问题、隐私保护问题、信任管理问题等。

物联网安全的研究应该突出从物联网应用中找安全需求,从有特色的共性网络技术中找安全问题,从物联网的特点中发现新问题。这里物联网的特点主要是指物联网存在多种形态网络的异构和融合、物联网设备可能具有资源受限的条件、设备可能是大规模且远距离可访问、设备的移动性和可定位追踪等。

从信息安全研究领域角度和信息安全需求角度,这里给出一个物联网安全的总体概貌,如图 1.2 所示。

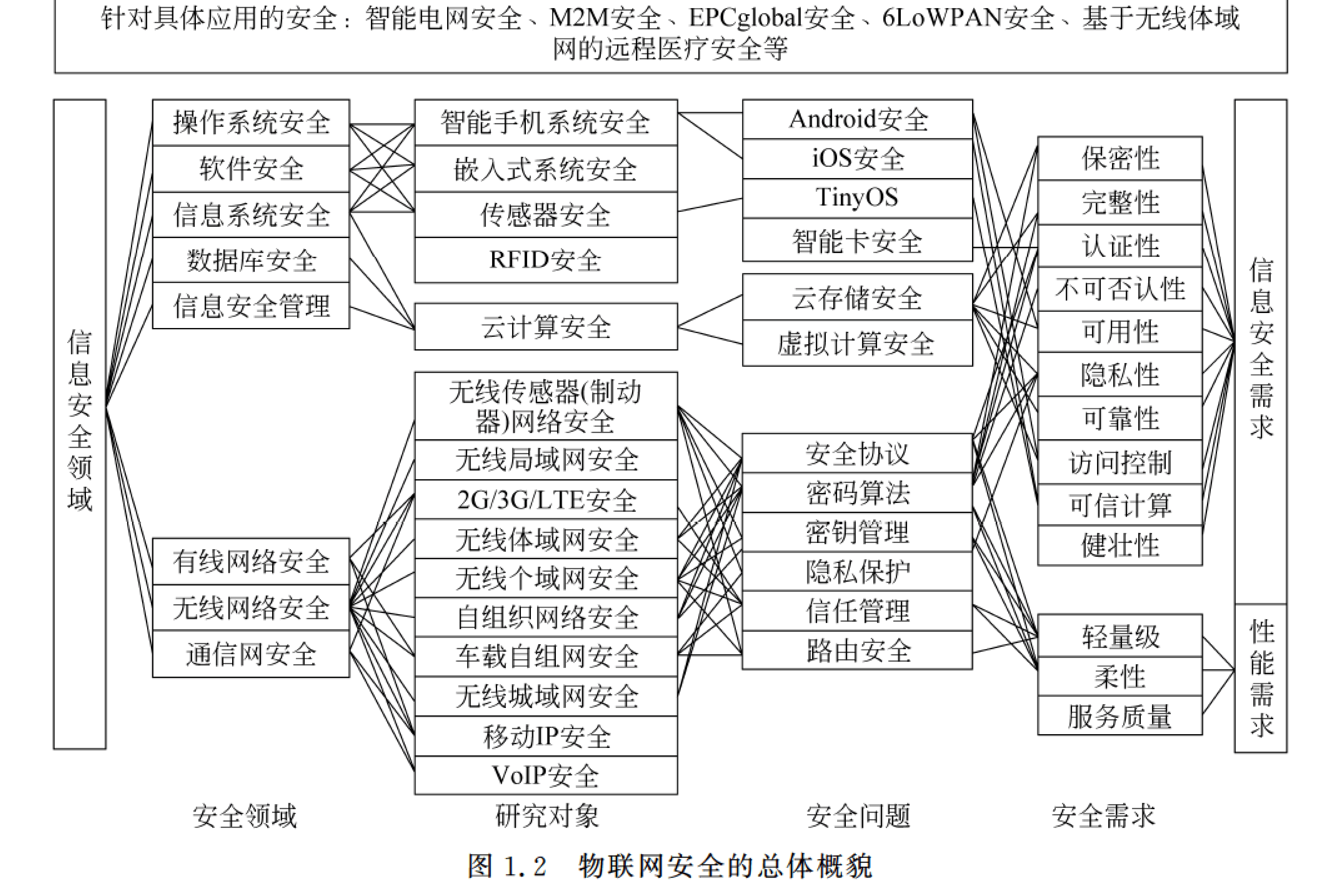


图 1.2　物联网安全的总体概貌

从物联网的架构出发，进行物联网安全的分类，可给出一个物联网安全架构的层次模型，如图 1.3 所示。这也是本书将要论述的主要内容，在每个论题中尽量选取了典型的网络情形和有代表性的安全问题。

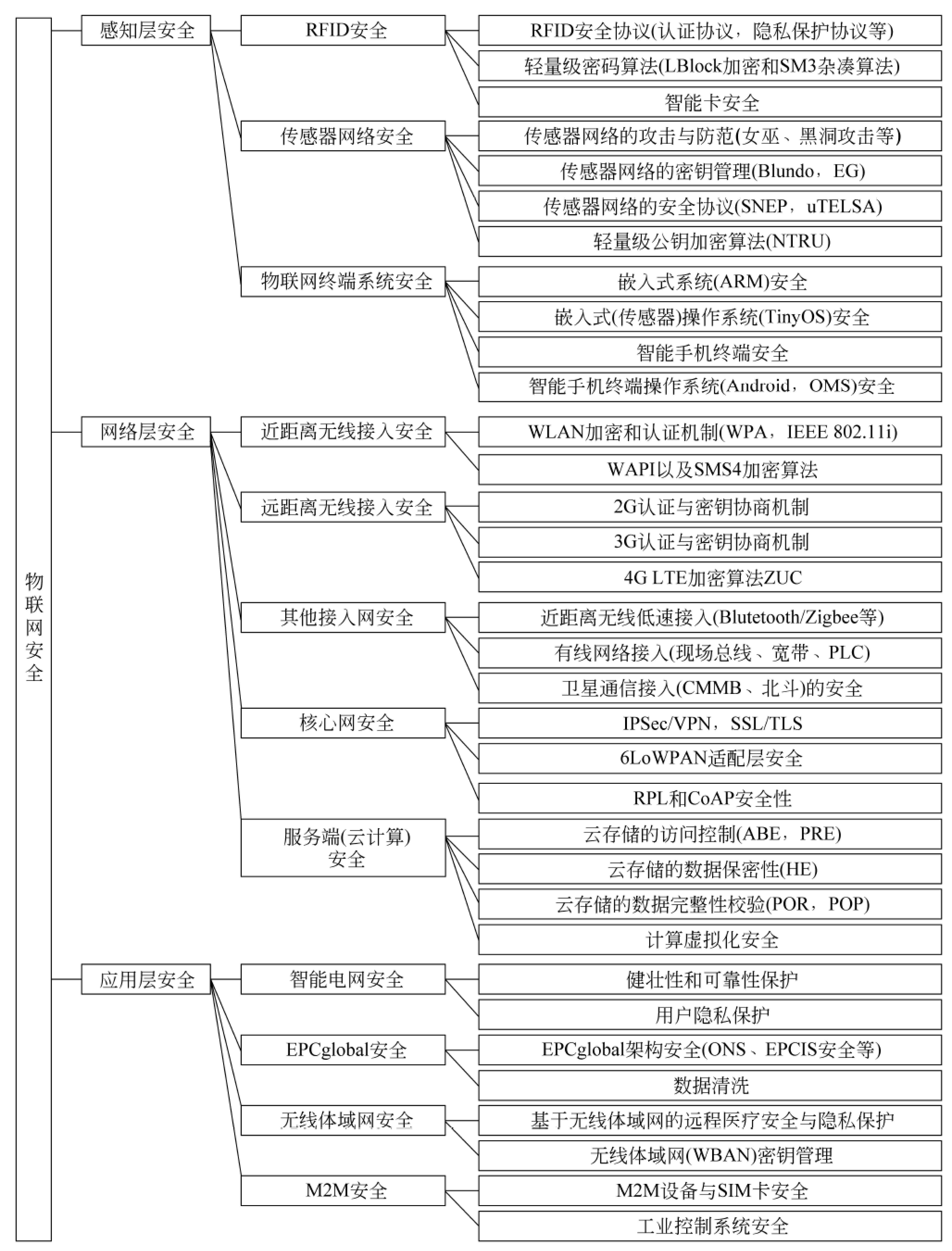


图 1.3　物联网安全架构的层次模型

下面给出一个物联网安全分析与设计的参考流程图，如图 1.4 所示。该流程图从感

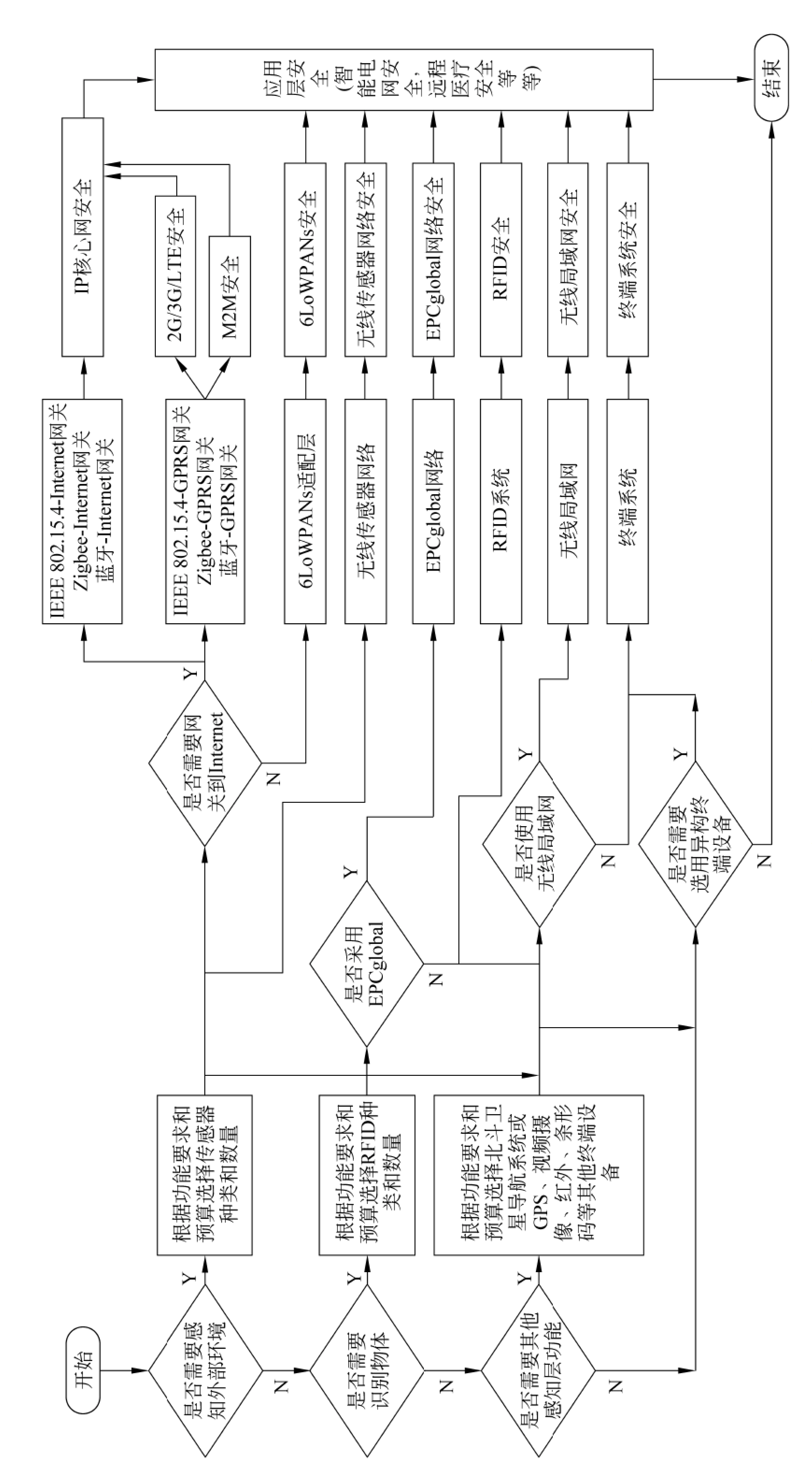


图 1.4　物联网安全设计的参考流程图

知层需求出发,根据网络架构的选择,确定相应的安全问题及其解决方案的范围。其中有些步骤不是严格区分的,但总体流程可作为实际中物联网安全分析和设计的参考。

1.2 网络安全问题的一般性讨论

1.2.1 物联网安全与相关学科的关联

为便于了解物联网安全的学科全貌,本节讨论物联网安全与相关学科的关联。

依据《信息安全专业指导性专业规范》,信息安全专业的研究内容可划分为4个领域:密码学、网络安全、信息系统安全和信息内容安全。每个领域的主干课程是:

(1) 密码学领域是密码学。

(2) 网络安全领域是网络安全、信息(网络)安全工程和信息(网络)安全管理。

(3) 信息系统安全领域是软件安全、操作系统安全、数据库安全、信息系统安全、电子商务和电子政务安全和智能卡技术。

(4) 内容安全领域是数字水印与信息隐藏,以及信息内容安全。

物联网安全应主要属于信息安全中的网络安全领域,特别是无线网络(含通信网)安全的范畴,但是,由于物联网的概念涵盖的范围非常广泛,例如包括了物联网的终端系统(如RFID、传感器结点、数据库系统以及服务器等),于是信息系统安全的内容如操作系统安全(如嵌入式操作系统、手机操作系统安全)、软件安全(智能手机病毒)、数据库安全等也会涉及。本书将重点讨论物联网安全的网络安全部分,包括无线网络(如大部分接入网的安全)和有线网络安全(如IP核心网安全),特别是具有物联网安全自身特色的内容,如6LoWPAN安全、智能电网、EPCglobal网络以及M2M,以区别一般的网络安全和无线网络安全。

通常网络安全的知识单元包括网络安全概念、防火墙、入侵检测系统(IDS)、虚拟专用网(VPN)、网络协议安全、网络安全漏洞检测与防护、Web安全等。这些多针对有线网络。无线网络安全的知识单元包括:无线局域网安全、无线城域网安全、无线广域网安全、无线个域网安全、无线体域网安全、无线自组织网络安全等。无线网络安全涵盖了通信网安全。物联网安全学科不是网络安全与无线网络安全研究内容的简单合并,而是在两者基础之上,更多地关注融合后新出现的安全问题(如6LoWPAN)以及新的网络形态下(如EPCglobal、M2M)的安全问题。

因此,在基本的如融合、异构、资源受限结点、大规模结点等约束条件下,或者在具体应用情形下如智能电网、M2M、远程医疗、控制网络等,或者特有网络架构下如6LoWPAN安全、EPCglobal网络等,去发现安全问题并解决这些问题时,这里更加强调利用密码学(特别是轻量级密码学)的方法,因为利用密码学这一解决信息安全问题的基本工具来解决物联网安全问题,可能会更加深刻、更加精巧。当然,也需要考虑到具体的安全需求,对于机密性、完整性、认证性、不可否认性问题,通常使用密码学工具。对于信任管理、隐私问题、可用性、健壮性等问题,则可以利用更多种类的方法。

图1.5在文献[2]的基础上进行了修改,标明了物联网安全课程在信息安全专业中的

位置。

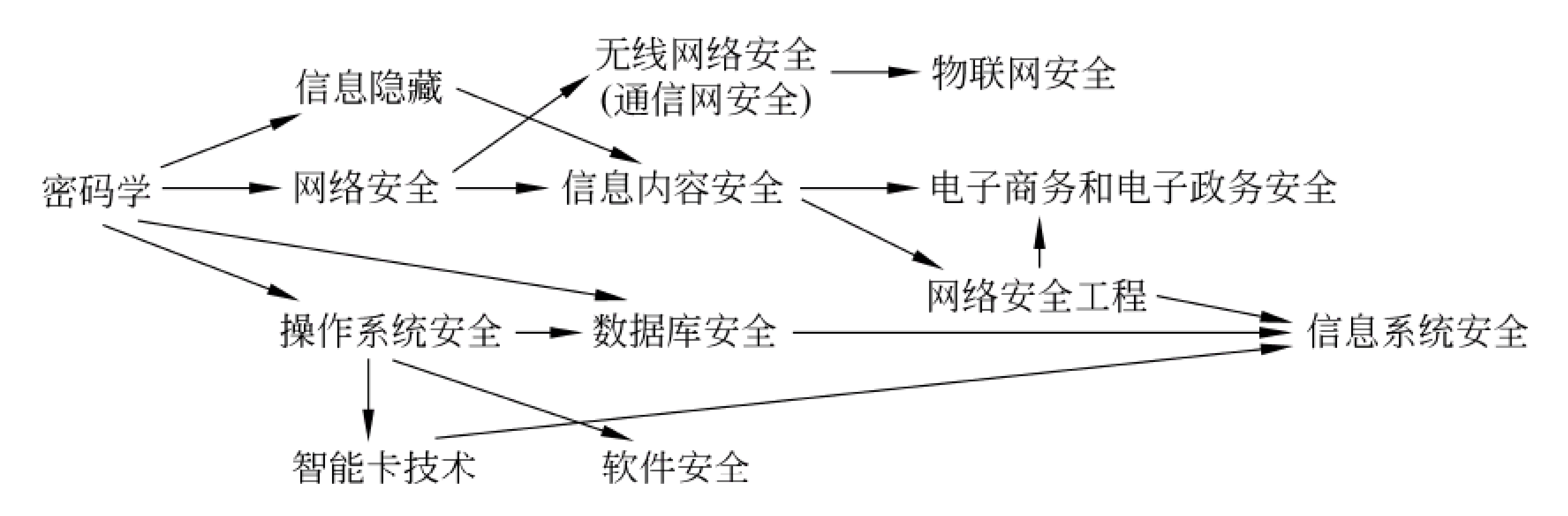


图 1.5　物联网安全在信息安全中的位置(侧重于网络安全方向)

这里简要回顾一下相关学科的全貌。

1. 信息安全

针对信息安全的攻击,主要包括主动攻击和被动攻击。被动攻击主要是信息的截取(Interception),指未授权地窃听传输的信息,企图分析出消息内容或者是通信模式。主动攻击包括:

(1) 中断(Interruption),阻止通信设施的正常工作,破坏可用性。

(2) 篡改(Modification),更改数据流。

(3) 伪造(Fabrication),将一个非法实体伪装成一个合法的实体。

(4) 重放(Replay)攻击,将一个数据单元截获后进行重传。

信息安全的目标通常包括:

(1) 机密性(Confidentiality)。指保证信息不泄漏给非授权的用户或者实体,确保保存的信息和被传输的信息仅能被授权的各方得到,而非授权用户即使得到信息也无法知晓信息的内容。通常通过访问控制机制阻止非授权的访问,通过加密机制阻止非授权用户或者信息的内容。

(2) 完整性(Integrity)。指消息未经授权不能进行篡改,要保证消息的一致性,即消息在生成、传输、存储和使用过程中不应发生人为或者非人为地非授权篡改(插入、修改、删除、重排序等)。一般通过访问控制阻止篡改行为,同时通过消息鉴别算法来检测信息是否被篡改。

(3) 认证性(Authentication)。指确保一个消息的来源或者消息本身被正确地标识,同时确保该标识没有被伪造,认证分为消息鉴别和实体认证。消息鉴别是指接收方保证消息确实来自于所声称的来源;实体认证指能确保被认证实体是所声称的实体,第三方不能假冒被认证的实体。

(4) 不可否认性(Non-Repudiation)。指能保证用户无法事后否认曾经对信息进行的生成、签发、接收等行为。当发送一个消息时,接收方能证实该消息确实是由既定的发送方发来的,称为源不可否认性;同样,当接收方收到一个消息时,发送方能够证实该消息确实已经送到了指定的接收方,称为宿不可否认性。一般通过数字签名来提供不可否认服务。

(5) 可用性(Availability)。指保障信息资源随时可提供服务的能力。即授权用户根

据需要可以随时访问所需信息，保证合法用户对信息资源的使用不被非法拒绝。典型的对可用性的攻击是拒绝服务攻击。

除了以上一些主要目标外，还有隐私性(Privacy)、匿名性(Anonymity)等。

为达到上述目标，信息安全采用了信息论、计算机科学和密码学等方面的知识，形成了一门综合学科，其主要任务是研究计算机系统和通信网络中信息的保护方法，以及实现系统内信息的机密性、完整性、认证性、不可否认性、可用性等，其中密码学是实现信息安全目标的核心技术。

2. 密码学

密码学(Cryptology)研究了实现信息安全各目标的相关的数学、方法和技术。密码学不是提供信息安全的唯一方式。其研究的目标是信息安全目标的一个子集，主要包括：机密性、完整性、认证性、不可否认性。为实现上述目标，密码学结合数学、计算机科学、电子与通信等诸多学科的方法于一体，是一门交叉学科。从大的方面可分为密码编码学和密码分析学两类，分别对应于密码方案的设计学科和密码方案的分析学科。

围绕着密码学要达到的目标，可以将密码学的实现方案分类成各种工具。图 1.6 给出了密码学内容的构成，并围绕着安全目标给出了各内容间的联系。

3. 物联网工程

物联网安全的研究是物联网研究中的重要组成部分，甚至是最关键的一个部分，一个不安全的物联网毫无疑问是危险的。物联网的研究将带动整个信息类学科如计算机、通信、自动控制、电子工程、电气工程等的研究。物联网的学术会议上可看到自不同专业领域和不同专业背景的研究人员。

国家设立物联网工程为一门新兴产业专业，是为了加快支撑物联网相关产业的专业人才的培养，实现可持续发展战略。作为工科的物联网工程专业内容涉及的专业课程来自包括计算机科学与工程、电子与电气工程、电子信息与通信、自动控制、遥感与遥测、精密仪器等专业中的部分课程。

物联网工程专业中与物联网安全相关的主干课程包括：物联网产业与技术导论、无线传感网络概论、TCP/IP 网络与协议、嵌入式系统技术、传感器技术概论、RFID 技术概论、工业信息化及现场总线技术、M2M 技术概论。从专业课程角度来讲，物联网安全是物联网工程专业的核心专业主干课之一。它同时也可作为信息安全、计算机科学等其他物联网相关专业的专业选修课。

1.2.2 一般性安全威胁及其具体表现

从信息安全角度来说，安全威胁是指某个人、物体或事件对某一资源的保密性、完整性、可用性或合法使用性所造成的危险。安全威胁可分为故意的和偶然的，故意的威胁又可进一步分为主动的和被动的。偶然的威胁是随机的，通常从可靠性和容错性角度进行分析；故意的威胁具有智能性，危害性更大，通常是安全分析中的主要内容。被动威胁只对信息进行监听，而不进行修改。主动威胁包括对信息进行故意篡改(包含插入、删减、添加等)、伪造虚假信息等。对每一种可能的攻击行为都要从攻击方法、攻击可能导致的后果、攻击者的数量与位置、实施这种攻击的可能性、攻击产生的先决条件和特征等方面进

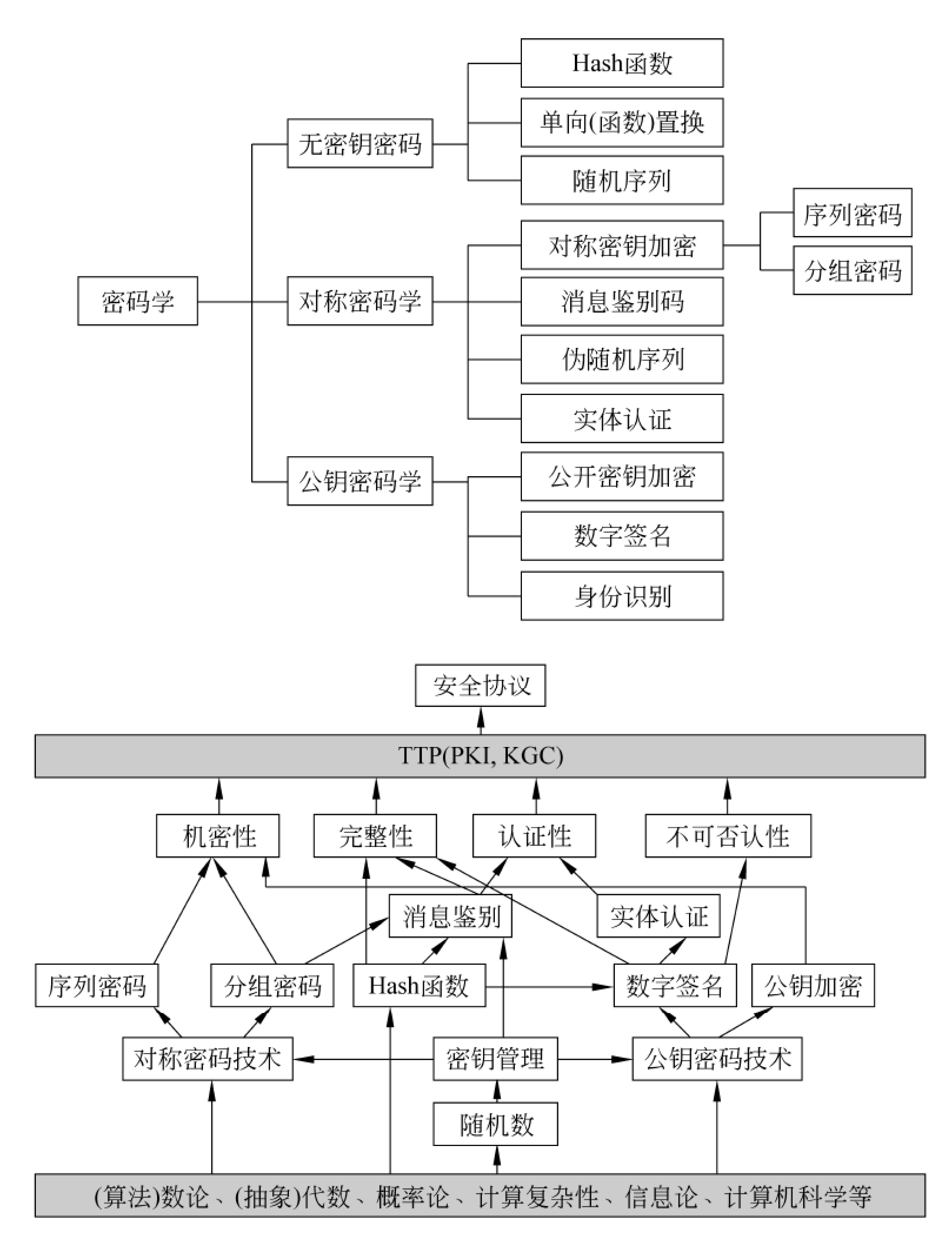


图 1.6　密码学的内容构成以及密码学研究内容间的关系

行分析，以便采取相应的安全对策。

通常，网络环境中安全威胁的具体表现主要有以下 3 个方面。

（1）无线以及有线链路上存在的安全威胁：通常在制定网络安全方案时，无线链路的安全威胁都需要得到充分的考虑，此外，与无线链路相连的有线链路（可能是骨干核心网）也需要同时加以考虑。考虑链路中的安全边界。由于先前一些无线网络中的有线链路部分可视为不开放的独立网络，其安全隐患往往被忽视，但随着无线网络的不断发展和互连，先前的有线骨干核心网络部分可能不再是孤立和封闭的，有线链路受到的威胁也需要加以考虑。具体表现如下。

① 攻击者被动窃听链路上的未加密信息，或者收集并分析使用弱密码体制加密的信息。

② 攻击者篡改、插入、添加或删除链路上的数据。攻击者重放截获的信息已达到欺

骗的目的。

③ 因链路被干扰或攻击而导致移动终端和无线网络的信息不同步或者服务中断。

④ 攻击者从链路上非法获取用户的隐私，包括定位、追踪合法用户的位置、记录用户使用过的服务、根据链路流量特征推测用户个人行为的隐私等。

（2）网络实体上存在的安全威胁：具体表现如下。

① 攻击者伪装成合法用户使用网络服务。攻击者伪装成合法网络实体欺骗用户使其接入，或者与其他网络实体进行通信，从而获取有效的用户信息，便于展开进一步攻击。

② 合法用户超越原有权限使用网络服务。

③ 攻击者针对无线网络实施阻塞、干扰等攻击。

④ 用户否认其使用过某种服务、资源或完成的某种行为。

（3）移动终端上存在的安全威胁：包括移动终端由于丢失或被窃取而造成其中的机密信息泄漏；现有移动终端操作系统缺乏完整性保护和完善的访问控制策略，容易被病毒或恶意代码所侵入，造成用户的机密信息被泄漏或篡改。

从信息安全的4个基本安全目标（机密性、完整性、认证性以及可用性）的角度来看，可将安全威胁相应地分成4大类基本威胁：信息泄漏、完整性破坏、非授权使用资源和拒绝服务攻击。围绕着这4大类主要威胁，在网络中具体的安全威胁主要有无授权访问、窃听、伪装、篡改、重放、重发路由信息、错误路由信息、删除应转发消息、网络泛洪(Flooding)等。表1.1给出了具体的安全威胁。

表1.1　具体安全威胁与基本安全威胁的对应关系

具体安全威胁	基本安全威胁			
	信息泄漏	完整性破坏	非授权使用资源	拒绝服务攻击
非授权访问	√		√	√
窃听	√			
伪装	√	√	√	√
篡改		√	√	√
重放			√	√
重路由、错误路由、删除信息		√	√	√
网络泛洪				√

从网络通信服务（通信网安全）的角度来看，安全防护措施通常称为安全业务，具体如表1.2所示。

表1.2　与安全威胁相对应的安全业务

安全威胁	安全业务					
	访问控制	实体认证	数据来源认证	数据完整性	数据机密性	不可否认性
非授权访问	√	√	√			
窃听					√	

续表

安全威胁	安全业务					
	访问控制	实体认证	数据来源认证	数据完整性	数据机密性	不可否认性
伪装		√	√	√		
篡改			√	√		
重放			√(时效性)			
重路由、错误路由、删除信息				√		
网络泛洪	√					
抵赖						√

表 1.2 的安全业务主要来自信息安全中的密码学角度,如果加上计算机网络安全的服务则还包括安全报警、安全响应和安全审计等安全服务。

*1.2.3 解决物联网网络安全问题的一般思路

物联网安全问题应该通过分析物联网的特点和特征入手,紧密结合具体实际应用分析安全需求,其讨论空间涵盖了无线网络安全和有线网络的安全,需要兼顾的知识面更为宽泛。由于本书的侧重点是物联网安全的网络安全部分,因此,很大程度上可以借鉴针对无线网络安全的研究方法。

在具体分析物联网安全问题时,可参考的一般思路如下(如图 1.7 所示)。

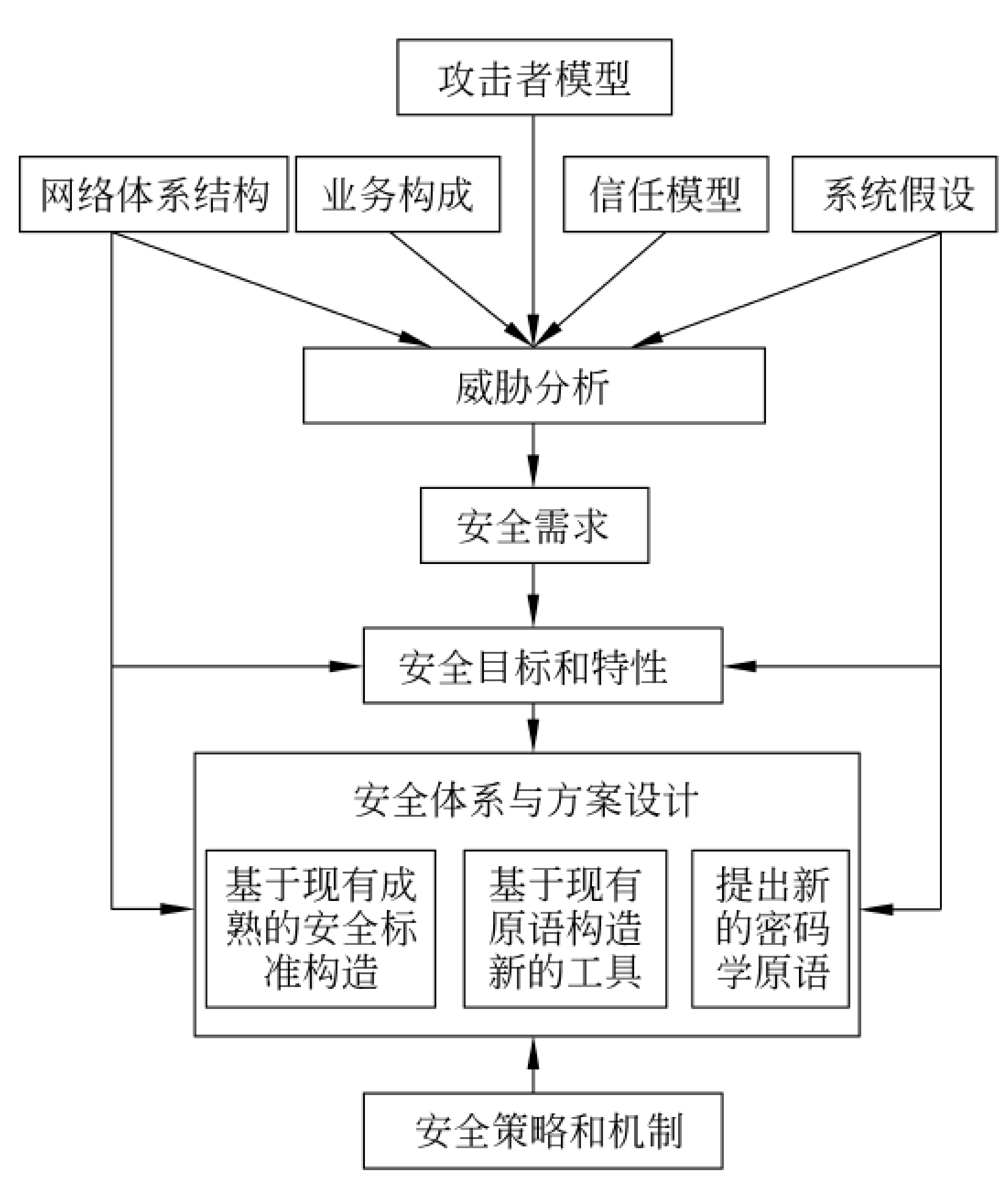


图 1.7 解决物联网安全问题的一般思路

（1）分析对系统的假设和约定。这主要指对网络终端、网络中间实体等网络结点系统的假设与约定，通常包括对网络中各相关结点的计算、通信、存储、电源等能力的假设。相同的安全问题，对于不同的假设和约定条件下，通常导致不同的解决方法。例如，网络终端结点的计算能力是否有限制，如 RFID 和传感器结点在计算能力上是有区别的，能够部署和执行的安全算法是有差异的，传感器结点上一般采用轻量级的密码算法，如 NTRU、TinyECC 等，RFID 上能够采用的加密算法多为轻量级分组算法，如 LBlock 等。

（2）分析网络的体系结构，明确网络的拓扑结构（星形、网状、分层树状、单跳还是多跳网络、拓扑结构是否变化、结点是否移动、结点移动的速度范围）、通信类型（单播、组播、广播等）、链路特征参数（带宽、吞吐率、延迟）、网络规模（结点数量、网络覆盖面积）、业务数据类型（语音、数据、多媒体、控制指令）等。以及网络的异构性（多种形态网络的融合，有线网络和无线网络的融合），网络的时效性（是临时存在的还是长期存在的）。它和上一条一起构成了设计安全方案时的客观约束条件，例如，网络拓扑结构往往会影响路由安全，结点移动性会影响身份认证，网络规模会影响密钥管理，业务数据类型会影响加密方式等。这些条件也会影响到后面对信任模型和攻击者模型的建模。例如，临时的动态的网络通常没有可信第三方，异构网络中的有线核心网部分是否存在攻击者。

（3）分析网络的业务构成（工作流程、操作过程），涉及的实体（角色）、业务通信的基本内容等，思考这些实体和通信内容可能面临的安全威胁。例如，网络的业务构成过程中可能遭受的安全威胁，业务的工作流程决定了需要安全保护的具体通信内容，涉及的实体决定了协议设计中的交互方以及访问控制对象。这一部分的分析将帮助确定具体的安全威胁，并最终帮助确定对应的安全需求。

（4）分析网络和系统中的信任模型，明确方案涉及的相关实体和通信链路的信任程度，即通信链路或者实体是可信、半可信还是不可信的，思考并确定安全的边界。信任模型中半可信的一个例子是指能够按照协议执行相关操作，但会泄漏或者篡改协议通信的内容。某些不可信的攻击者可能不按照所期望网络协议的方式操作，如无线传感器网络中的 Blackhole 攻击、Greyhole 攻击、Wormhole 攻击、Sybil 攻击等，这时需要借助非密码学的方法，如入侵检测、基于信任的管理等机制等。

（5）分析攻击网络和系统的攻击者模型：是内部还是外部攻击，是主动还是被动攻击，思考对攻击者能力的设定（固定攻击者还是移动攻击者），给出一些典型的攻击场景，以及对这些攻击可能导致的后果。例如，在无线传感器网络中特殊的攻击方式 Sybil 攻击、虫洞攻击，RFID 网络中的隐私破坏问题，针对网络编码的 Pollution 攻击等等。如果对攻击者模型的假设越强，则安全性越高。根据网络的特征来分析，便于发现该网络中存在的特有的安全威胁或攻击模式，防御这些威胁时，通用的网络安全措施可能不能奏效，这是便需要根据该网络的业务特点以及相关系统和体系结构的假设与约定进行安全方案设计。发现新的攻击方法是无线网络安全研究中的一个基本创新点，其创新之处在于发现并提出了一个新的安全问题。如果进而给出对新的攻击方法的安全方案则构成了一个完整的创新点。

（6）从存在的威胁中归纳出共性的安全需求。通常的思路是从信息安全基本安全需求的角度来分析，包括保密性、认证性、完整性、可用性、健壮性（鲁棒性、容侵、容错、抗结

点妥协、可靠性)、隐私保护、信任管理。无线网络的移动性特点和设备的不可靠性特点,使得隐私保护和健壮性这两种安全需求更加受到重视。安全需求一般是与具体的安全威胁相对应,也可能是将安全威胁进行归纳后的涵盖安全威胁的最小集合。根据现代密码学的要求,安全需要通过形式化的方法进行严格的定义,这种定义往往会用到可忽略函数,概率多项式图灵机,概率不可区分性等基本概念。

(7) 根据前面步骤中归纳的安全需求、网络体系结构、系统假设确定设计需要达到的安全目标,以及实现该目标的时要满足的特性,例如安全算法需要满足的计算量上限,存储空间上限,安全方案对容侵容错的健壮性等。思考在满足网络体系结构和系统假想条件下如何满足完全需求。思考安全防御的总体思路,例如是采用密码学的方法,还是采用与通信网络和计算机安全相关的方法,如人工智能的方法、概率统计的方法、信任评价和管理的方法、博弈论的方法等。

(8) 根据安全目标和特性、网络体系结构、系统假设等最后确定安全体系或方案。如果有成熟的密码学机制(如安全标准、安全算法和安全协议),则通常是一种根据实际应用背景的工程设计。根据实际应用背景对相关密码学机制进行修改和应用,这一设计的需要考虑的要点是:对安全标准技术的工程应用选择、信息安全技术应用在实际场合的合理性、必要性、完备性,以及对历史遗留系统的兼容性,部署安全方案的成本代价。安全策略和机制则更多地从网络管理和安全管理的角度考虑实际中安全方案的性能和可用性。

(9) 实际应用的需要可能需要构造新的组合工具。根据现有密码学原语构造适用于应用环境的组合工具(也称为复杂系统),其特点是利用现有的(抽象)基本密码学原语或者基本工具,如伪随机生成器、伪随机函数、抗碰撞单向函数、陷门置换函数等构造复杂的组合工具、协议、安全方案等。这一种设计往往根据设计者的个人经验,有高度的创造性,甚至具有一定的偶然性。因此,如何根据安全目标进行逻辑推理,进而构造自动的安全方案设计工具,乃至进行自动的安全性证明,笔者认为将是一个非常有前途的研究方向。

(10) 在某些特殊的应用场景中,现有密码学原语及其组合工具可能不能满足安全需求,这时可以思考是否需要提出新的密码学原语(或者是密码学构造的新需求),例如签密,以及某种新的特殊数字签名方案(不可否认签名、聚集签名、在线离线签名、批量验证签名等)、具有特殊属性的公钥加密机制(代理重加密、基于属性的加密、完全同态加密等)、轻量级对称密钥加密机制、轻量级安全协议等。实践证明,新的密码学原语(或者构造方法)往往是来自于实际中的需求。虽然可能只是给出抽象的构造方法,没有给出实际的构造方案,但提出原语本身就具有一定创新性。无线网络安全研究中不断地丰富和完善实际构造方案是非常常见的。

研究与思考

[1] 给出一个你所想象的 2020 年物联网的应用场景,并思考这一场景中将会涉及哪些安全问题。

[2] 现在有些媒体报道了物联网保险柜、物联网热水器等。物联网保险柜就是当有人碰触保险柜时,保险柜的主人将收到短信通知。物联网热水器是指你可以发送短信给热水器开始烧热水等你下班回家。思考一下你如何来实现这些想法(提示:保险柜可以安装传感器,感知的判断可以通过

M2M 通信方式经过移动网络运营商的网络发送到你的手机上)。

[3] 在回家的路上,通过手机上的语音搜索你找到了一瓶 30 年的好葡萄酒,并知道在你家附近的商店就有这瓶酒。于是你通过手机支付下了订单,等你到家,葡萄酒已经送到家里,你通过手机上的 RFID 读写器识别这个葡萄酒是否是正品。思考一下,如何实现这一应用(提示:需要用到 EPCglobal 网络、定位系统方面的知识)。

进一步阅读建议

搜索并阅读一下关于物联网安全方面的综述文章。

[1] A. Shabtai, Y. Fledel, U. Kanonov, Y. Elovici, S. Dolev, C. Glezer, Google Android: A Comprehensive Security Assessment [J], IEEE Security & Privacy, 8(2):35-55, 2010.

[2] Internet of Things Strategic Research Roadmap[OL], http://www. internet-of-things-research. eu/pdf/IoT_Cluster_Strategic_Research_Agenda_2011. pdf.

[3] R. H. Weber, Internet of Things - New Security and Privacy Challenges, Computer Law and Security Report, Elsevier, 26(1): 23-30, 2010.

[4] 杨庚,许建,陈伟,祁正华,王海勇. 物联网安全特征与关键技术[J]. 南京邮电大学学报(自然科学版),2010, 30(4): 20-29.

[5] 沈苏彬,范曲立,宗平,毛燕琴,黄维. 物联网的体系结构与相关技术研究[J]. 南京邮电大学学报(自然科学版), 29(6): 1-11, 2009.

本章参考文献

[1] 刘云浩. 物联网导论[M]. 北京:科学出版社, 2011.

[2] 杜瑞颖,张焕国,王丽娜,陈晶. 本科信息安全专业课程体系研究[J],计算机教育,2011.8.

[3] 金纯,郑武,陈林星. 无线网络安全——技术与策略[M]. 北京:电子工业出版社, 2004.

[4] 姜楠,王健译. 移动网络安全技术与应用[M]. 北京:电子工业出版社, 2004.

[5] 张帆. 无线网络安全协议的形式化分析方法[D]. 西安电子科技大学博士学位论文, 2007.

[6] 朱建明. 无线网络安全方法与技术研究[D]. 西安电子科技大学博士学位论文, 2004.

[7] 任伟. 无线网络安全问题初探[J]. 信息网络安全, 2012.01.

[8] Wei Ren, QoS-aware and Compromise Resilient Key Management Scheme for Heterogeneous Wireless Internet of Things, International Journal of Network Management (Wiley), 2011, 21, 284～299.

[9] 任伟,物联网安全架构与技术路线研究[J]. 信息网络安全,2012.4.

[10] 任伟,公钥密码体制的基本原理研究[J]. 信息网络安全,2011.5.

[11] 任伟,密码学和现代密码学研究[J]. 信息网络安全,2011.8.

[12] 任伟,可证明安全公钥密码学探究[J]. 信息网络安全,2011.11.

第 1 部分　物联网感知层安全

第 2 章　RFID 安全

第 3 章　无线传感器网络安全

第 4 章　物联网终端系统安全

第 2 章　RFID 安全

2005 年，国际电信联盟（ITU）发表了题为《ITU 互联网报告 2005：物联网》的报告，该报告认为物联网主要有 4 个关键技术：RFID 技术、无线传感器网络技术、智能技术、纳米技术。其中前 3 个是信息技术，而前两个是与感知层有关的。

RFID 技术是物联网感知层的关键技术，主要功能是对物体的识别。其实，如果这里的物体是广义的，即也包括人、字符、图像等，则识别技术还可包括：光学符号识别技术（如对文字、数字的识别）、图像识别、语音识别技术、生物计量（Biometrics）识别技术（如虹膜识别和指纹识别）。当然，对物体的识别还可以是非 RFID 技术，例如磁卡技术、IC 卡（智能卡）技术、条形码技术（一维条形码、二维条形码）。磁卡和 IC 卡在银行卡中有广泛的应用。一维条形码在超市结算的商品识别中有广泛的应用，二维条形码则在移动互联网的网上“团购”网站中开始大量使用。表 2.1 比较了几种常见的物体识别技术。

表 2.1　RFID 标签与条形码、磁卡、IC 卡的性能比较

	信息载体	信息量	读/写	读取方式	保密性	智能化	抗干扰能力	寿命	成本
条形码	纸张、塑料薄膜、金属表面	小	只读	激光扫描	差	无	差	较短	最低
磁卡	磁性物质	一般	读/写	电磁转换	一般	无	较差	短	低
IC 卡	EEPROM	大	读/写	电擦除、写入	最好	有	好	长	较高
标签	EEPROM	大	读/写	无线通信	最好	有	很好	最长	较低

由于 RFID 具有典型性和重要性，因此本章重点介绍 RFID 的安全。

2.1　RFID 系统简介

2.1.1　RFID 系统的基本构成

无线射频识别（Radio Frequency Identification，RFID）是一种非接触式自动识别技术。它通过无线射频方式自动识别标签所附着的目标对象，获取 RFID 标签的相关信息。RFID 技术可识别高速运动的目标对象（例如不停车收费 ETC），并可同时识别多个标签，能够快速地进行物品的追踪和管理，具有可靠性高、保密性强、成本低廉等特点。它广泛应用于仓储管理、物品追踪、防伪、物流配送、过程控制、访问控制、门禁、自动付费、供应链管理、图书管理等领域。

RFID 系统一般由 3 部分组成（如图 2.1 所示）：标签（Tag）、读写器（Reader）以及后

端数据库(Back-end Database)。

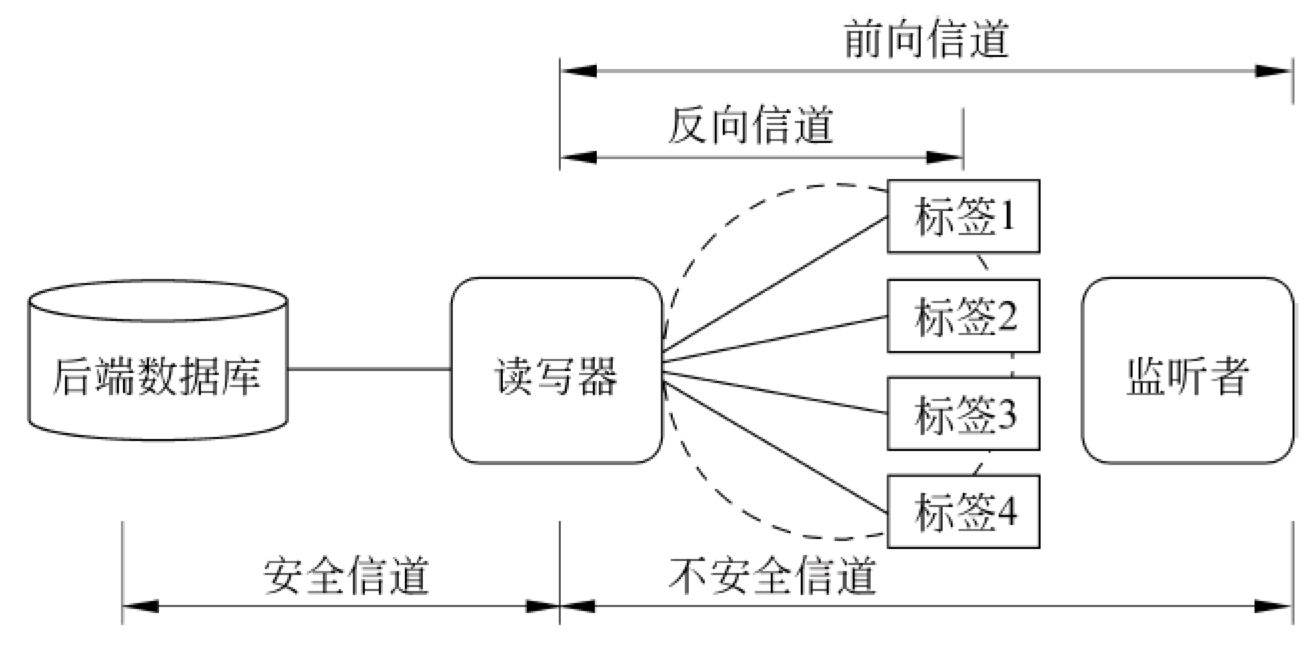


图 2.1　RFID 系统的基本构成

1. 标签

标签由专用芯片和天线或线圈组成，通过电感耦合或者电磁反射原理与读写器进行通信。它附着在目标对象上，如护照、身份证、人体、动物、物品、票据、手机(手机钱包应用)，用于识别或者跟踪目标对象。存储在标签芯片中的数据(即对物品的编码)用于唯一地标识被识别或者跟踪的对象。

根据标签的能量来源，通常将标签分为 3 大类：被动式标签(Passive Tag)、主动式标签(Active Tag)和半被动式标签(Semipassive Tag)。

被动式标签内部不带电池，需要靠外界提供能量才能正常工作，即需要从读写器发出的无线电波中获取能量。因而它能永久使用，常用于标签信息需要频繁读/写的场合，而且它支持长时间数据传输和长期数据存储。被动式标签比主动式标签轻便、便宜、寿命长，但它的传送距离短(第一代标签 1 米，第二代标签 3～5 米)且需要更高功率的读写器，灵敏度和定位性能受限于读写器，且存储数据的容量和抗噪声能力有限。

主动标签的能量来自内部电池，也称为有源标签，因此一个密封的主动标签寿命是有限的。和被动标签相比，主动标签传送距离更远(可达上百米)、速度更快、抗噪声更好，使用相同频率时，数据传输速率更高，但体积大、价格高。

半被动标签兼具被动式标签和主动式标签的优点。本身虽带有电池，但只起到对标签内部数字电路供电的作用。标签并不通过自身能量主动发送数据，只有当它被读写器的能量场“激活”时，才传送自身的数据。

2. 读写器

读写器实际上是一个带有天线的无线发射与接收设备，它的处理能力、存储空间都比较大。读写器分为手持和固定两种。RFID 标签是非接触的，借助读写器来实现数据的读/写功能，对 RFID 标签写入信息或者读取标签所携带的数据信息。

读写器到标签之间的信道叫做前向信道(Forward Channel)，而标签到读写器之间的信道叫做反向信道(Backward Channel)。由于读写器与标签的无线功率差别很大，前向信道的通信范围远大于反向信道的通信范围。

3. 后端数据库

后端数据库是一个数据库系统，通常假设其计算和存储能力强大，并包含所有标签的

信息。在独立的 RFID 应用环境中，通常假设标签和读写器之间的通信信道是不安全的，而读写器和后端数据库之间的通信信道则是安全的(在使用 RFID 的供应链应用或者在基于 EPCglobal 的物联网应用环境中，这一假设可能被改变)。

RFID 的基本工作原理如下：读写器发射电磁波，而此电磁波有其辐射范围，当标签进入此电磁波辐射范围内，标签将读写器所发射的微小电磁波能量存储起来，并转换成电路所需的电能，并且将存储的标识信息以电磁波的方式传送给读写器。标签和读写器之间的通信距离受到多个参数(特别是通信频率)的影响。

4. 典型研究对象

目前使用最多的第二代被动式标签。目前我国已经在第一代被动式标签即高频应用领域占据了世界第一的位置，形成了从芯片设计、制造、封装和读写器设计、制造到应用的成熟产业链(例如上海优比科公司的 RFID 天线/电子标签年产 30 亿张，年产能居全球首位)。目前国际上重点发展的是第二代被动式标签，即超高频领域，因此，这里对后者加以重点介绍。

第二代被动式标签采用超高频，范围为 300MHz～3GHz，3GHz 以上为微波范围，典型的工作频率为 433MHz，860～960MHz，2.45GHz，5.8GHz。通信距离一般大于 1 米，典型为 4～6 米，最大可超过 10 米。超高频阅读器有非常高的数据传输速率，在很短的时间内可读取大量标签。超高频的电子标签数据存储量一般限定在 2048 位以内。EPCglobal 规定的 EPC 容量为 96 位。典型应用包括供应链管理、生产线自动化、航空(铁路)包裹管理、集装箱管理等。

第一代被动式标签采用高频，范围为 3～30MHz，通信距离一般小于 1 米，价格相对便宜。典型应用包括图书管理系统、酒店门锁管理、(医药、服装)物流系统、智能货架管理等，两者的比较见表 2.2。

表 2.2　第二代和第一代被动式标签的比较

比　　较	超 高 频	高　　频	
协议	EPC Gen2 (ISO 18000—6C)	ISO 15693	ISO 14443
频率(MHz)	860-960(区域选择)	13.56(全球统一)	
通信距离(m)	3～5	1	0.1
存储大小(位)	96～1000	256～64K	
读写器价格(美元)	500～2000	100～1000	
标签价格(美元)	0.1～0.2	0.2～0.5	
应用	供应链、自动化、资产管理与跟踪	访问控制、安全付款、验证	

EPCglobal 网络安全将在第 11 章介绍。

2.1.2　RFID 系统的安全需求

1. 安全问题

简单而言，RFID 的安全问题包括隐私问题和认证问题。隐私问题是由读写器读标

签时无须认证引起的，包括信息泄漏问题和追踪问题；认证问题是由标签被读取时无须认证引起的，包括标签克隆、篡改标签数据等。

广义的 RFID 系统的隐私保护包括两点：一是标签和读写器之间的隐私保护；另一种是服务器中的信息隐私保护，它所关心的是服务器所包含什么样的信息。本小节主要讨论第一种情况。

隐私问题中的信息泄漏问题是指在获取标签信息之后可对 RFID 系统进行各种非授权使用，标签可泄漏相关物体和用户信息，如护照、身份证、贵重物品标签持有者可能成为抢劫或者盗窃的目标，个人的药品信息被他人所知导致个人的隐私泄漏等。

隐私问题中的追踪问题是指通过标签的唯一标识符可恶意地追踪用户的位置或者行为。例如，标识在不同的地方的两次出现，说明用户曾经到达了这两个地方。攻击者可在任何地点、任何时间追踪识别某个固定标签，从而侵犯了用户隐私。

隐私问题可通过对读写器的认证来解决。认证问题可通过对标签的认证来解决。因此，读写器和标签之间的双向认证是 RFID 系统的主要安全需求。

2. 安全威胁

具体而言，一个安全的 RFID 系统应该对以下攻击加以防范。

(1) 非法读取：非法者通过未授权的读写器读取标签中的数据信息。

(2) 窃听：标签和读写器之间的数据传输容易受到窃听攻击。

(3) 无前向安全性：攻击者在此次通信中截取到了标签的输出，然后通过某种推算可以得出标签以前所发送的信息。反过来说，如果攻击者不能推算前面发出的信息，则称为前向安全性(Forward Security)。

(4) 位置跟踪：非法者通过标签发出的固定消息来定位标签的位置以达到跟踪的目的。

(5) 伪装：非法者截取到标签信息后，把真实标签信息复制到自己假冒的标签中。当读写器发送认证消息给标签时，非法者把自己复制的标签信息发给读写器，以伪装成合法标签通过读写器的认证。

(6) 重放：当读写器发出认证信息时，攻击者截取了标签发出的响应信息。当下一次读写器发出认证请求时，攻击者把截取到的信息发送给读写器，从而通过读写器对它的认证。

(7) 拒绝服务攻击：许多基于挑战-应答方式的协议都要求每次对标签进行访问时，标签都需要提供额外的存储器来存储要产生的随机数，当大量读写器向标签发送询问信息时，标签的存储器就因要存储过多的随机数而停止工作。

3. 安全方案设计时的考虑因素

为了设计 RFID 的安全方案，需要考虑到 RFID 标签的计算能力，这种计算能力通常限定了可采用的安全方案。一般可把 RFID 标签的计算能力分为以下 3 类。

(1) 基本标签：不能执行加密操作，但可执行 XOR 操作和简单的逻辑控制的标签。

(2) 对称密码标签：指能够执行对称密钥加密操作的标签。

(3) 公钥密码标签：指能够执行公钥加密操作的标签。

2.2　RFID 安全的物理机制

实现 RFID 安全性机制所采用的方法主要有 3 大类：物理机制、密码机制，以及二者相结合的方法。本节介绍的物理机制主要有如下几类：Kill 命令机制、休眠机制、阻塞机制、静电屏蔽、主动干扰等。物理机制通常用于一些低成本的标签中，因为这些标签难以采用复杂的密码机制来实现与标签读写器之间的安全通信。下面分别加以介绍。

Kill 命令机制：由 Auto-ID Center 提出的 Kill 命令机制是解决信息泄漏的一个最简单的方法。即从物理上毁坏标签，一旦对标签实施了 Kill 命令，标签便不能再次使用（禁用状态）。例如，超市结账时可禁用附着在商品上的标签。但是，如果 RFID 标签用于标识图书馆中的书籍，当书籍离开图书馆后，这些标签是不能被禁用的，这是因为当书籍归还后还需要使用相同的标签再次标识书籍。

休眠（Sleeping）机制：让标签处于睡眠状态，而不是禁用，以后可使用唤醒口令将其唤醒。困难在于唤醒口令需要和标签相关联，于是这就需要一个口令管理系统。但是，当标签处于睡眠状态时，不可能直接使用空中接口将特定的标签和特定的唤醒口令相关联。因此需要另一种识别技术，例如条形码，以标识用于唤醒的标签，这显然是不理想的。

阻塞（Blocking）机制：隐私位 0 表示标签接受无限制的公共扫描；隐私位 1 表示标签是私有的。当商品生产出来，并在购买之前，即在仓库、运输汽车、存储货价的时候，标签的隐私位置为 0。换句话说，任何读写器都可扫描它们。当消费者购买了使用 RFID 标签的商品时，销售终端设备将隐私位设置为 1。

法拉第网罩：也称为静电屏蔽法。由于无线电波可被传导材料做成的电容屏蔽，将贴有 RFID 标签的商品放入由金属网罩或金属箔片组成的容器中，从而阻止标签和读写器通信。由于每件商品都需使用一个网罩，提高了成本。

主动干扰：标签用户通过一个设备主动广播无线电信号用于阻止或破坏附近的 RFID 读写器操作。但该方法可能干扰附近其他合法 RFID 系统，甚至阻塞附近其他无线电信号系统。

2.3　RFID 安全密码协议

RFID 安全的密码机制指利用各种成熟的密码方案和机制来设计和实现符合 RFID 安全需求的密码协议。现有的基于密码技术的 RFID 安全机制大致可以分为两大类：“静态 ID 机制”及“动态 ID 刷新机制”。所谓“静态 ID 机制”就是标签的标识保持不变，而“动态 ID 刷新机制”则是标签的标识随着每一次标签与读写器之间的交互而动态变化。采用动态 ID 刷新机制时，一个非常重要的问题就是“数据同步问题”，也就是说，后端数据库中所保存的标签标识必须和存储在标签中的标识同步进行刷新，否则，在下一次认证识别过程中就可能使得合法的标签无法通过认证和识别。另外，RFID 协议在设计时需要注意的地方是对 RFID 标签的计算能力的假定。

目前已经提出了大量RFID安全协议,这里重点介绍几个典型的认证协议,如Sarma等人的Hash锁协议[1]、Weis等人的随机化Hash锁协议[2]、Ohkubo等人的Hash链协议[3]、X. Gao等人的Good Reader协议[4]、D. Molnar等人的David数字图书馆协议[5]等。

2.3.1 Hash锁协议

Hash锁(Hash-Lock)协议使用metaID来代替真实的标签ID。它是一种基于单向Hash函数的机制,每一个具有Hash锁的标签中都有一个Hash函数并存储一个临时metaID。具有Hash锁的标签可以工作在锁定和非锁定两种状态,锁定状态下的标签,对所有问询的响应仅仅是metaID;标签只有在非锁定状态时才向邻近的读写器提供它的信息。后面的描述中用H来表示单向Hash函数。

1. 标签的锁定过程

标签的锁定过程如下:

(1) 读写器选定一个随机密钥key,并计算metaID=H(key)。

(2) 读写器写metaID到标签。

(3) 标签进入锁定状态。

(4) 读写器以metaID为索引,将(metaID,key,ID)存储到本地后端数据库。

2. Hash-Lock协议的执行过程

Hash-Lock协议的执行过程如图2.2所示。

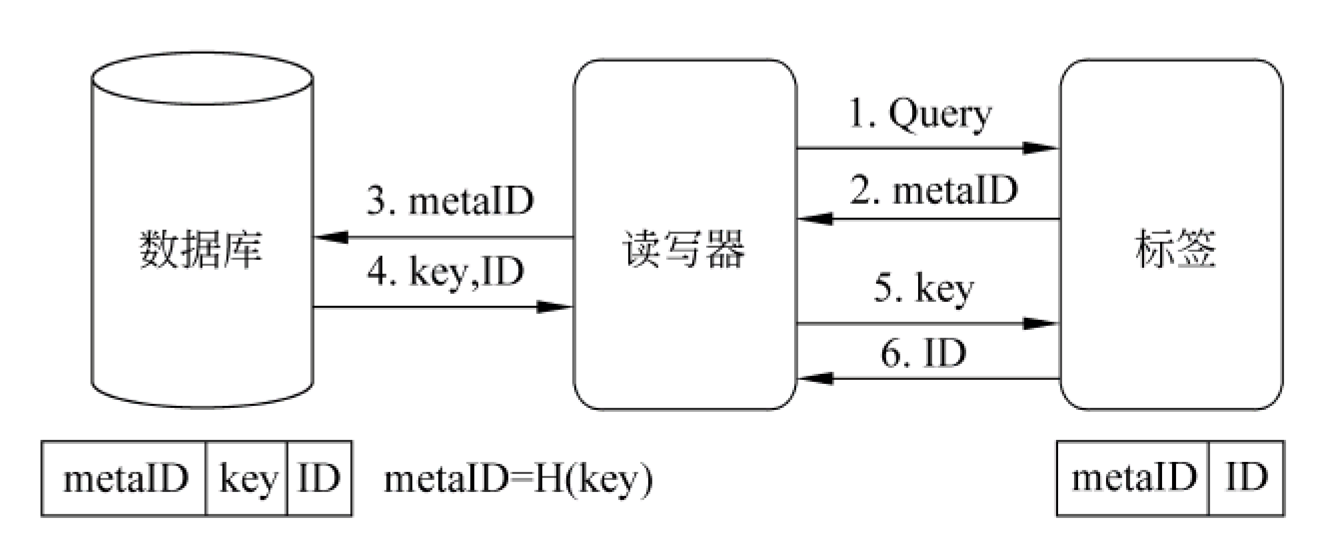


图 2.2 Hash-Lock协议

(1) 读写器向标签发送认证请求,用Query表示,即向标签发出问询其标识的请求。

(2) 标签将metaID发送给读写器。

(3) 读写器将metaID转发给后端数据库。

(4) 后端数据库查询自己的数据库,如果找到与metaID匹配的项,则将该项的(key,ID)发送给读写器,其中ID为待认证标签的标识;否则,返回给读写器认证失败信息。

(5) 读写器将从后端数据库接收的部分信息key发送给标签。

(6) 标签验证metaID=H(key)是否成立,如果成立,则对读写器的认证通过,将其ID发送给读写器;否则认证失败。

(7) 读写器比较从标签接收到的ID是否与后端数据库发送过来的ID一致,若一致,则对标签的认证通过;否则认证失败。

Hash-Lock协议虽然是双向认证,但没有ID动态刷新机制,且metaID也保持不变,

ID 以明文的形式通过不安全的信道传送，因此该协议非常容易受到假冒攻击和重传攻击，所谓假冒攻击就是利用窃听到的 metaID 和 ID 伪造一个标签代替真实的标签，通常能通过读写器的认证。重传攻击的方法也是类似。另外，因为 ID 是不变的，所以攻击者可以很容易地对标签进行追踪。因此 Hash-Lock 协议没有达到保护 ID 不泄漏的安全目的。

2.3.2　随机化 Hash 锁协议

随机化 Hash 锁协议是 Hash 锁协议的改进。对原来的 metaID 进行了随机化，使其总是变化的，从而试图避免可追踪性。随机化 Hash 锁协议采用了基于随机数的“挑战——应答”机制，即认证方提问，被认证方回答，如果回答正确，则说明被认证方通过了认证方的认证。其协议流程如图 2.3 所示。

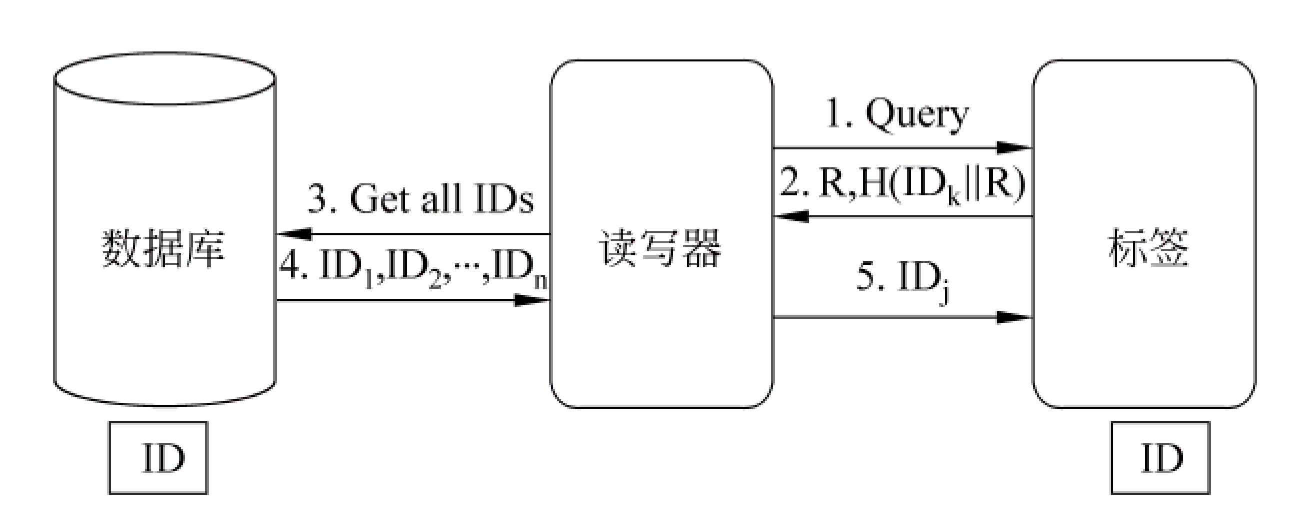


图 2.3　随机化 Hash 锁协议

随机化 Hash 锁协议的执行过程如下：

(1) 读写器向标签发送认证请求 Query。

(2) 标签生成一个随机数 R，计算 $H(ID_k \| R)$，其中 ID_k 为标签的标识。标签将 $(R, H(ID_k \| R))$ 发送给读写器。

(3) 读写器向后端数据库请求获得所有标签的标识。

(4) 后端数据库将自己数据库中的所有标签的标识 $(ID_1, ID_2, \cdots, ID_n)$ 发送给读写器。

(5) 读写器检查是否有某个 $ID_j (1 \leqslant j \leqslant n)$，使得 $H(ID_j \| R)$ 成立；如果有，则对标签的认证通过，并且把 ID_j 发送给标签。

(6) 标签验证。ID_j 与 ID_k 是否相同，如相同，则对读写器的认证通过。

该认证协议也是双向认证，虽然消息 2 中是不断变化的，但该认证过程仍然存在问题：认证通过后的标签标识 ID_j 仍以明文的形式在不安全信道传送，因此攻击者还是可以对标签进行有效地追踪。同时，一旦获得了标签的标识 ID_j，攻击者就可以对标签进行伪造。因此，随机化 Hash 锁协议也是不安全的。另外，每一次标签认证时，后端数据都需要将所有标签的标识发送给读写器，二者之间的数据通信量很大，效率也就很低。

2.3.3　Hash 链协议

Hash 链协议是基于共享秘密的“挑战-应答”协议。在该协议中，当不同读写器发起

认证请求时，若两个读写器中的 Hash 函数不同，则标签的应答就是不同的。其协议流程如图 2.4 所示。

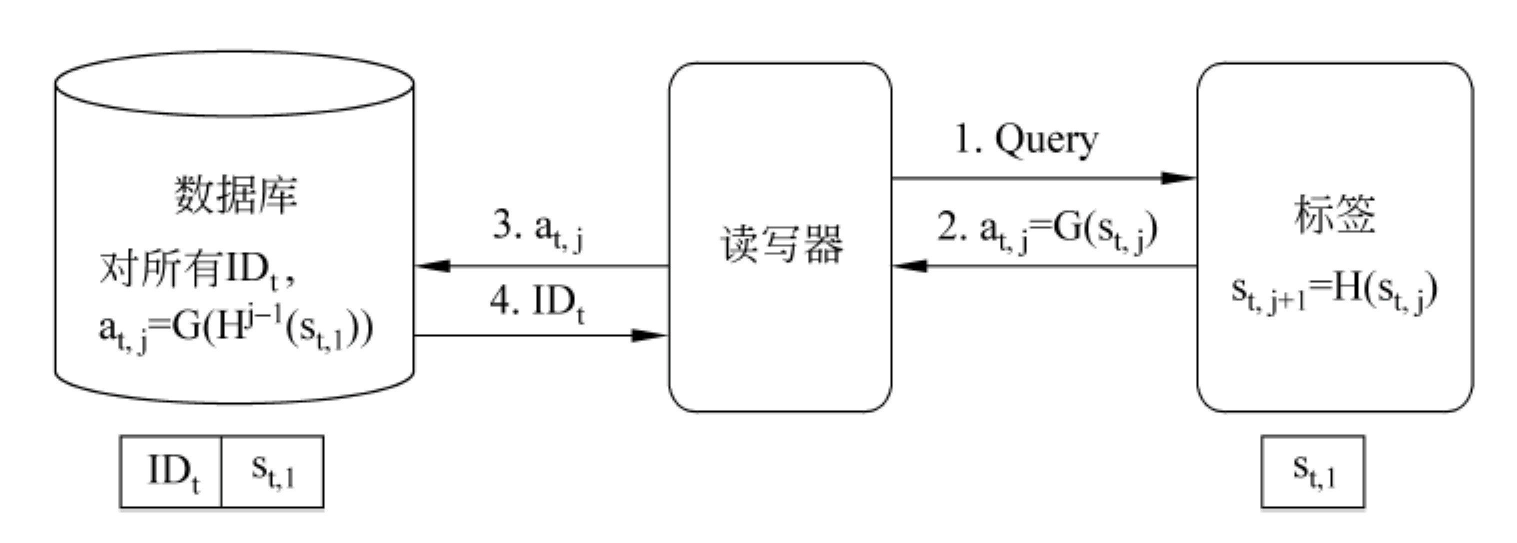


图 2.4　Hash 链协议

在系统运行之前，标签和后端数据库首先要共享一个初始秘密值 $s_{t,1}$，则标签和读写器之间执行第 j 次 Hash 链的过程如下：

（1）读写器向标签发送认证请求 Query。

（2）标签使用当前的秘密值 $s_{t,j}$ 计算 $a_{t,j}=G(s_{t,j})$，并更新其秘密值为 $s_{t,j+1}=H(s_{t,j})$，标签将 $a_{t,j}$ 发送给读写器（G 也是一个密码学安全的 Hash 函数）。

（3）读写器将 $a_{t,j}$ 转发给后端数据库。

（4）后端数据库系统针对所有的标签数据项查找并计算是否存在某个 $ID_t(1\leqslant t\leqslant n)$ 及是否存在某个 $j(1\leqslant j\leqslant m)$，其中 m 为系统预先设定的最大链长度，使得 $a_{t,j}=G(H^{j-1}(s_{t,1}))$ 成立。如果有，则认证通过，并将 ID_t 发送给标签；否则，认证失败。

该协议满足了不可追踪性，因为 G 是单向函数，攻击者观察到的 $a_{t,j}$ 和 $a_{t,j+1}$ 是不可关联的。另外满足前向安全性，即使攻击者从窃听的 $a_{t,j}$ 来推算出了 $s_{t,j}$，也无法知道 $s_{t,j-1}$ 等，从而无法知道 $a_{t,j-1}$，还是无法进行追踪。

上述协议仍然容易受到重传和假冒攻击，只要攻击者截获某个 $a_{t,j}$，它就可以进行重传攻击，伪装成合法标签通过认证。此外，每一次标签认证发生时，后端数据库都要对每一个标签进行 j 次 Hash 运算，计算量相当大。同时，该协议需要至少两个不同的 Hash 函数，增加了实现标签功能需要的电路门的数量，从而增加了成本。

2.3.4　Good Reader 协议

Good Reader 协议是单向认证协议，认证读写器是否合法，需要约定读写器标识 ReaderID 首先被存放在标签中，标签可以通过它所存储的读写器标识来验证读写器的合法性，该协议的认证过程如图 2.5 所示。

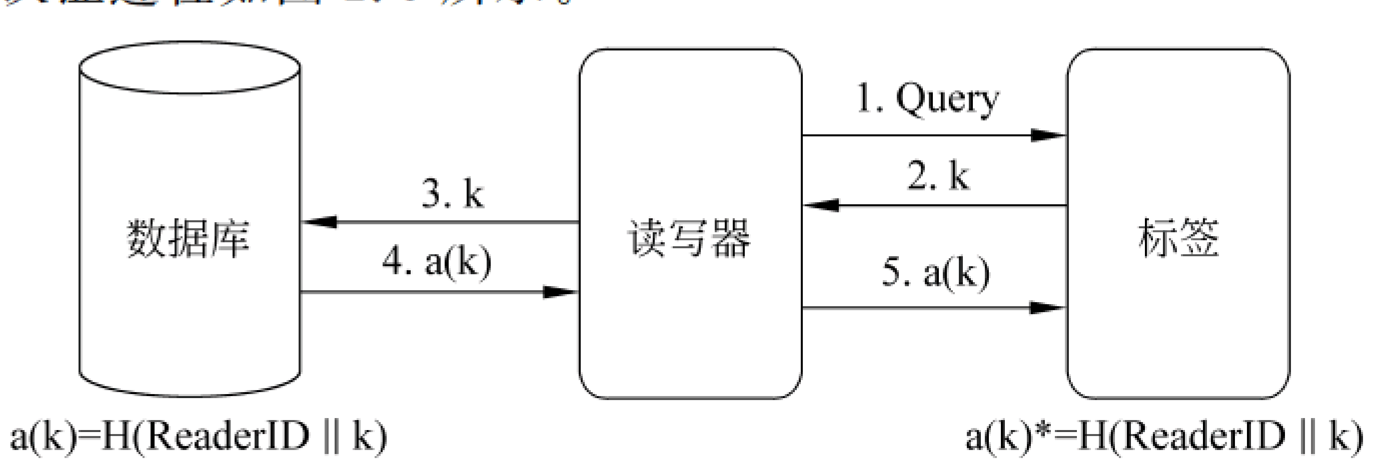


图 2.5　Good Reader 协议

协议过程如下：

(1) 当读写器发送问询消息给标签以开始认证。

(2) 标签产生一个随机数，并把此随机数发送给读写器，同时，标签计算 $a(k)^* = H(ReaderID \| k)$。

(3) 读写器将接收到的随机数发送给后台数据库，后台数据库利用 k 和已存有的读写器标识 ReaderID 进行计算：$a(k) = H(ReaderID \| k)$。然后，数据库将所得数据 $a(k)$ 发送给读写器。

(4) 读写器将接收到的 $a(k)$ 发送给标签。

(5) 标签通过比较先前计算过的 $a(k)^*$ 和接收到的 $a(k)$ 是否相等来判别读写器的合法性。

此协议可以有效防止因固定输出而引起的位置跟踪和假冒的问题，并且数据库中不需要进行大量的 Hash 运算，这样大大缩短了运算时间。但该协议没有对标签进行认证，且标签中需要存储读写器的标识，大大增加了标签的存储容量，从而增加了标签的成本。

2.3.5 David 数字图书馆协议

David 数字图书馆 RFID 协议使用基于预共享秘密的伪随机函数来实现认证，其协议流程如图 2.6 所示。系统运行之前，后端数据库和每一个标签之间需要预先共享一个秘密值 s，该协议的执行过程如下：

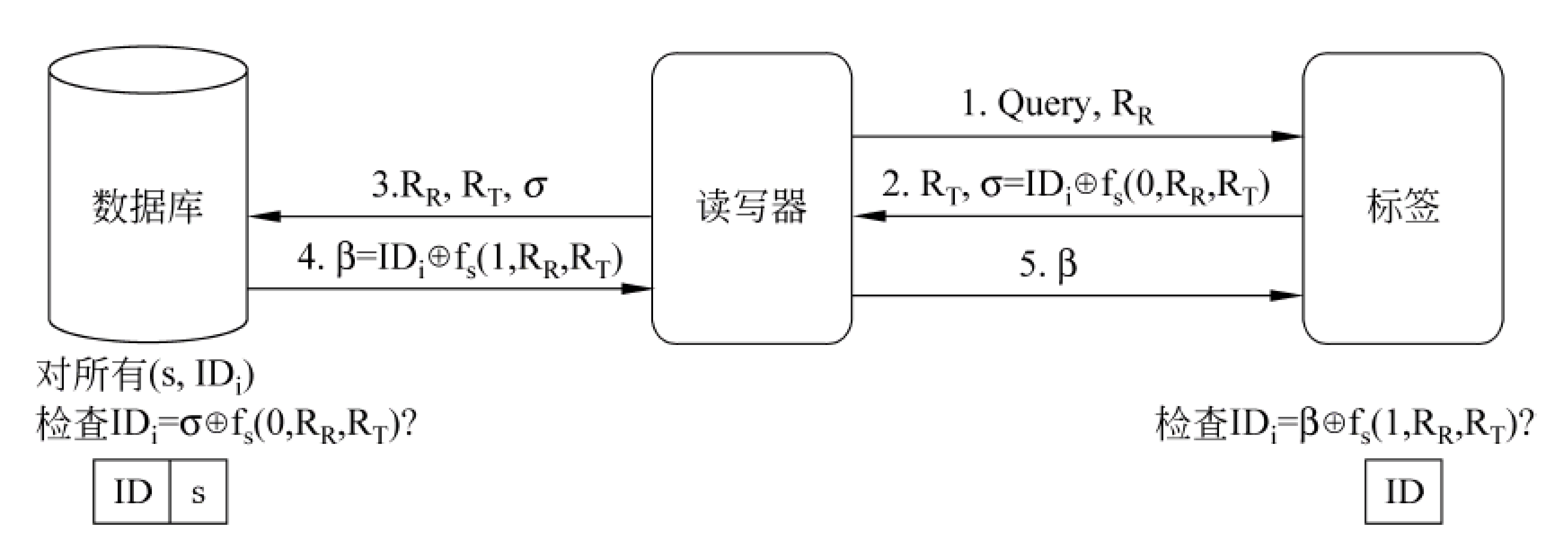


图 2.6　David 数字图书馆 RFID 协议

(1) 读写器生成一个秘密随机数 R_R，向标签发送认证请求 Query，将 R_R 也发送给标签。

(2) 标签生成一个随机数 R_T，使用自己的 ID 和秘密值 s 计算 $\sigma = ID_i \oplus f_s(0, R_R, R_T)$，标签将 (R_R, R_T, σ) 发送给读写器。

(3) 读写器将 (R_R, R_T, σ) 转发给后端数据库。

(4) 后端数据库检查是否有某个 $ID_i (1 \leqslant i \leqslant n)$，满足 $ID_i = \sigma \oplus f_s(0, R_R, R_T)$；如果有，则对标签的认证通过，并计算 $\beta = ID_i \oplus f_s(1, R_R, R_T)$，然后将 β 发送给读写器。

(5) 读写器将 β 转发给标签。

(6) 标签验证 $ID_i = \beta \oplus f_s(1, R_R, R_T)$ 是否成立，如成立，则对读写器的认证通过。

该协议必须在标签电路中包含实现随机数生成器和安全伪随机函数两大功能模块，故而该协议不适用于低成本的 RFID 系统。

*2.4 密码算法

2.4.1 轻量级分组加密算法 LBlock

在 RFID 这样的资源受限环境中实现的分组加密必须是轻量级的,轻量级密码处理的数据规模较小,因而加密算法的数据吞吐量的要求比普通分组加密要低得多。另外,轻量级分组密码大多采用硬件实现,要求实现占用的空间要小,这个通常用算法实现需要的等效门数 GE(Gate Equivalent)来衡量。GE 表示独立于制造技术的数字电路复杂性的度量单位。设计轻量级分组加密的方法主要有两种:在现有算法基础上,对密码算法的组件进行轻量级改进;一种是设计一个全新的轻量级算法。目前比较知名的轻量级分组加密有 PRESENT、HIGHT、CGEN、DESL、MIBS 等。中国学者吴文玲设计了一种轻量级分组密码[6],中文名叫“鲁班锁”,英文名叫 LBlock,既是 Luban lock 的缩写,也有 Lightweight Blockcipher 的意思。

LBlock 的分组长度为 64 位,密钥长度为 80 位。算法对差分密码分析、线性密码分析、不可能差分密码分析、相关密钥攻击等具有足够的安全冗余。算法具有优良的硬件实现效率,仅仅需要 866.3GE,同时在 8 位和 32 位处理器上有很好的实现性能。下面简要介绍一下该算法。

1. 加密算法

加密算法由 32 轮迭代运算组成,对 64 位的明文 $P=X_1 \| X_0$,加密过程如下:

对 i = 2,3, …, 33,计算

$$X_i=F(X_{i-1},K_{i-1})\oplus(X_{i-2}<<<8)$$

$X_{32} \| X_{33}=$ C 为 64 位的密文

其中的基本模块定义如下所示。

(1) 轮函数 F。

$$F:\{0,1\}^{32}\times\{0,1\}^{32}\rightarrow\{0,1\}^{32}$$

$$(X,K_i)\rightarrow U=P(S(X\oplus K_i))$$

(2) 函数 S。

函数 S 是函数 F 的一部分,由 8 个 4×4 的 S 盒并置而成,定义如下:

$$S:\{0,1\}^{32}\rightarrow\{0,1\}^{32}$$

$$Y=Y_7\|Y_6\|Y_5\|Y_4\|Y_3\|Y_2\|Y_1\|Y_0\rightarrow Z=Z_7\|Z_6\|Z_5\|Z_4\|Z_3\|Z_2\|Z_1\|Z_0$$

$$Z_7=S_7(Y_7),\ Z_6=S_6(Y_6),\ Z_5=S_5(Y_5),\ Z_4=S_4(Y_4)$$

$$Z_3=S_3(Y_3),\ Z_2=S_2(Y_2),\ Z_1=S_1(Y_1),\ Z_0=S_0(Y_0)$$

(3) 函数 P。

函数 P 是 8 个 4 位字的位置变换,定义如下:

P：$\{0,1\}^{32} \to \{0,1\}^{32}$
$Z=Z_7 \parallel Z_6 \parallel Z_5 \parallel Z_4 \parallel Z_3 \parallel Z_2 \parallel Z_1 \parallel Z_0 \to U=U_7 \parallel U_6 \parallel U_5 \parallel U_4 \parallel U_3 \parallel U_2 \parallel U_1 \parallel U_0$
$U_7=Z_6$，$U_6=Z_4$，$U_5=Z_7$，$U_4=Z_5$
$U_3=Z_2$，$U_2=Z_0$，$U_1=Z_3$，$U_0=Z_1$

2. 解密算法

解密算法是加密算法的逆运算，有 32 轮迭代运算组成。对 64 位的密文 $C=X_{32} \parallel X_{33}$，解密过程如下：

对 j=31,30,…,1,0，计算
$X_j=F(X_{j+1},K_{j+1}) \oplus X_{j+2} >>> 8$
$X_1 \parallel X_0=P$ 为 64 位的明文

3. LBlock 的密钥扩展算法

将密钥 $K=k_{79}k_{78}k_{77}k_{76}\cdots k_1k_0$ 放在 80 位的寄存器中，取寄存器左边的 32 位作为轮密钥 K_1，然后执行如下步骤：

对 i=1,2,3,…,31，按如下方式更新寄存器
(1) $K <<< 29$
(2) $[k_{79}k_{78}k_{77}k_{76}] \leftarrow S_9 \leftarrow [k_{79}k_{78}k_{77}k_{76}]$
$[k_{75}k_{74}k_{73}k_{72}] \leftarrow S_8 \leftarrow [k_{75}k_{74}k_{73}k_{72}]$
(3) $[k_{50}k_{49}k_{48}k_{47}k_{46}] \oplus [i]_2$
(4) 取寄存器左边的 32 位作为轮密钥 K_{i+1}

其中，S_0 到 S_9 是 10 个 4×4 的 S 盒。

S 盒的详细设计以及 LBlock 的设计原理参见文献[7]。

2.4.2　密码 Hash 算法 SM3

国家密码管理局 2010 年 12 月公布了密码杂凑算法（Cryptographic Hash Algorithm）SM3[8]，包括计算方法和计算步骤，并给出了运算示例。该算法适用于商用密码应用中的数字签名和验证、消息认证码的生成与验证以及随机数的生成，可满足多种密码应用的安全需求。

1. 符号约定

算法描述中用到的符号见表 2.3。

表 2.3　SM3 算法中用到符号约定

ABCDEFGH：8 个字寄存器或它们的值的串联（字长是 32 位）
$B^{(i)}$：第 i 个消息分组
CF：压缩函数
FF_j：布尔函数，随 j 的变化取不同的表达式
GG_j：布尔函数，随 j 的变化取不同的表达式
IV：初始值，用于确定压缩函数寄存器的初态
P_0：压缩函数中的置换函数

续表

P_1：消息扩展中的置换函数 T_j：常量，随 j 的变化取不同的值 m：消息 m′：填充后的消息 mod：模运算 ∧：32 位与运算 ∨：32 位或运算 ⊕：32 位异或运算 ¬：32 位非运算 +：mod 2^{32} 算术加运算 <<<k：循环左移 k 位运算 ←：左向赋值运算符

2. 常数与函数

1）初始值

$$IV = 7380166f\ 4914b2b9\ 172442d7\ da8a0600\ a96f30bc\ 163138aa\ e38dee4d\ b0fb0e4e$$

2）常量

$$T_j = \begin{cases} 79cc4519 & 0 \leqslant j \leqslant 15 \\ 7a879d8a & 16 \leqslant j \leqslant 63 \end{cases}$$

3）布尔函数

$$FF_j(X,Y,Z) = \begin{cases} X \oplus Y \oplus Z & 0 \leqslant j \leqslant 15 \\ (X \wedge Y) \vee (X \wedge Z) \vee (Y \wedge Z) & 16 \leqslant j \leqslant 63 \end{cases}$$

$$GG_j(X,Y,Z) = \begin{cases} X \oplus Y \oplus Z & 0 \leqslant j \leqslant 15 \\ (X \wedge Y) \vee (\neg X \wedge Z) & 16 \leqslant j \leqslant 63 \end{cases}$$

式中 X，Y，Z 为字。

4）置换函数

$$P_0(X) = X \oplus (X <<< 9) \oplus (X <<< 17)$$
$$P_1(X) = X \oplus (X <<< 15) \oplus (X <<< 23)$$

式中 X 为字。

3. 算法描述

1）概述

对长度为 $l(l<2^{64})$位的消息 m，SM3 杂凑算法经过填充和迭代压缩，生成杂凑值，杂凑值长度为 256 位。

2）填充

假设消息 m 的长度为 l 位。首先将位 1 添加到消息的末尾，再添加 k 个 0，k 是满足 $l+1+k \equiv 448 \bmod 512$ 的最小的非负整数。然后再添加一个 64 位比特串，该比特串是长度 l 的二进制表示。填充后的消息 m′的位长度为 512 的倍数。

例如，对消息 01100001 01100010 01100011，其长度 l=24，经填充得到比特串：

$$01100001\ 01100010\ 01100011\ 1\ \overbrace{00\cdots00}^{423位}\overbrace{00\cdots011000}^{64位}$$

（后 64 位为 l的二进制表示）

3）迭代压缩

（1）迭代过程。

将填充后的消息 m′按 512 位进行分组：

$$m'=B^{(0)}B^{(1)}\cdots B^{(n-1)}$$

其中 n=(l+k+65)/512。

对 m′按下列方式迭代：

```
FOR i=0 TO n-1
    V(i+1) =CF(V(i) ,B(i) )
ENDFOR
```

其中 CF 是压缩函数，$V^{(0)}$ 为 256 位初始值 IV，$B^{(i)}$ 为填充后的消息分组，迭代压缩的结果为 $V^{(n)}$。

（2）消息扩展。

将消息分组 $B^{(i)}$ 按以下方法扩展生成 132 个字 $W_0,W_1,\cdots,W_{67},W'_0,W'_1,\cdots,W'_{63}$，用于压缩函数 CF：

```
将消息分组 B(i) 划分为 16 个字 W0,W1,…,W15
FOR j=16 TO 67
Wj←P1(Wj-16⊕Wj-9⊕(Wj-3<<<15))⊕(Wj-13<<<7)⊕Wj-6
ENDFOR
FOR j=0 TO 63
W'j=Wj⊕Wj+4
ENDFOR
```

（3）压缩函数。

令 A,B,C,D,E,F,G,H 为字寄存器，SS1,SS2,TT1,TT2 为中间变量，压缩函数 $V^{i+1}=CF(V^{(i)};B^{(i)})$，$0\leqslant i\leqslant n-1$。计算过程描述如下：

```
ABCDEFGH←V(i)
FOR j=0 TO 63
SS1←((A<<<12) + E + (Tj<<<j))<<<7
SS2←SS1⊕(A<<<12)
TT1←FFj(A;B;C) + D + SS2 +W'j
TT2←GGj(E;F;G) + H + SS1 +Wj
D←C
C←B<<<9
B←A
A←TT1
H←G
```

```
G←F <<<19
F←E
E←P0(TT2)
ENDFOR
V(i+1)←ABCDEFGH⊕V(i)
```

其中，字的存储为大端(big-endian)格式。所谓大端格式是指数据在内存中的一种表示格式，规定左边为高有效位，右边为低有效位。数的高阶字节放在存储器的低地址，数的低阶字节放在存储器的高地址。

4) 输出杂凑值

$ABCDEFGH \leftarrow V^{(n)}$

输出 256 位的杂凑值 y=ABCDEFGH。

研究与思考

[1] (使用 FPGA 等方法)实现 LBlock 算法并给出性能分析。

[2] (使用 C 语言)实现 SM3 算法并进行随机碰撞测试。

[3] 本节介绍主要针对包括读写器、标签和后端数据库的 RFID 系统。在没有后端数据库支持的情况下，或者后端数据库是可能攻陷的情况下，新的安全协议应该如何设计。

[4] 在物联网环境下，RFID 的一个最常见的应用就是物流管理，这往往需要引入一个管理中心，监控从不同地点的读写器读到的标签数据。在这种情形下，如何分析新的安全威胁并给出相应的安全协议设计。

进一步阅读建议

RFID 的安全协议设计是当前物联网安全的研究热点之一。关于 RFID 安全的更新较快且所列文献非常齐备的文献列表是 RFID Security & Privacy Lounge[1]。传统 RFID 系统协议和物联网供应链协议是有区别的，文献[2]给出了一个可证明安全的 RFID 供应链协议。基于 EPCglobal 的物联网安全将在第 11 章介绍。

[1] RFID Security & Privacy Lounge[OL]，http://www.avoine.net/rfid.

[2] 张帆，孙璇，马建峰，曹春杰，朱建明. 供应链环境下通用可组合安全的 RFID 通信协议[J]. 计算机学报，2008.10.

本章参考文献

[1] S. Sarma，S. Weis，D. Engels，RFID Systems And Security And Privacy Implications[C]，In Proc. of the 4th International Workshop on Cryptographic Hardware and Embedded Systems (CHES02). Berlin：Springer，2002：454-469.

[2] S. Weis，Security and Privacy In Radio Frequency Identification Device [M]，Cambridge，MA，

USA：MIT，2003.
[3] M. Ohkubo，K. Suzuki，S. Kinoshita，Hash-chain based Forward-Secure Privacy Protection Scheme For Lowcost RFID[C]，In Proc. of the 2004 Symposium on Cryptography and Information Security (SCIS04)，Sendai，2004，719-724.
[4] X. Gao，Zhe. Xiang，Hao Wang et a1.，An Approach to Security and Privacy of RFID System for Supply Chain[C]，In Proc. of the IEEE International Conference on E-Commerce Technology for Dynamic E-Business，Beijing，China，2004：164-168.
[5] D. Molnar，D. Wagner，Privacy and Security in Library RFID：Issues，Practices，and Architectures[C]，In Proc. of the 11th ACM Conference on Computer and Communications Security(CCS04)，Washington，DC，USA，2004，210-219.
[6] 中国密码学会组编，中国密码学发展报告 2010[M].北京：电子工业出版社，2011.
[7] 吴文玲，张蕾等.鲁班锁轻量级分组密码详细设计. 信息安全国家重点实验室技术报告，2010.9.
[8] SM3密码杂凑算法，国家密码管理局，http://www.oscca.gov.cn/UpFile/20101222141857786.pdf.
[9] 周永彬，冯登国. RFID 安全协议的设计与分析[J]. 计算机学报，2006，29(04)：581-589.
[10] 曹天杰，张永平，汪楚娇. 安全协议[M]. 北京：北京邮电大学出版社，2009.
[11] 刘宇靓，任伟.基于通用可组合理论的协议安全性证明方法讨论. 信息网络安全，2012.05.

第 3 章　无线传感器网络安全

正如第 2 章开头所介绍的，根据 ITU 的物联网报告，无线传感器网络是物联网的第二个关键技术。RFID 的主要功能是对物体的识别；而无线传感器网络的主要功能是感知。无线传感器网络则是大范围多位置的感知。通俗地说，传感器是可以感知外部环境参数的小型计算结点，传感器网络是大量传感器结点构成的网络，用于不同地点、不同种类的参数的感知或数据的采集，无线传感器网络则是利用无线通信技术来传递感知的数据的网络。

其实感知技术还可包括更多的方面，如红外线技术可以感知物体对光线的“遮挡”，广泛应用于节水龙头；摄像头可以感知（采集）物体的图像和动作，广泛应用于视频监控；GPS 设备可以感知物体的位置；声音感应控制电灯的开关。由于无线传感器网络是感知技术中最重要的一种，本章重点介绍无线传感器网络的安全。

3.1　无线传感器安全简介

无线传感器网络（Wireless Sensor Networks，WSN）是集成了传感器技术、微机电系统技术、无线通信技术以及分布式信息处理技术于一体的新型网络。随着科学技术的发展，信息的获取变得更加纷繁复杂。所有保存事物状态、过程和结果的物理量都可以用信息来描述。传感器的发明和应用，极大地提高了人类获取信息的能力。传感器信息获取从单一化到集成化、微型化，进而实现智能化、网络化，成为获取信息的一个重要手段。无线传感器网络在很多场合（如军事感知战场、环境监控、道路交通监控、勘探、医疗等）都承担重要的作用。

3.1.1　无线传感器网络的体系结构

1. 传感器结点的物理结构

在不同的应用场景中，传感器结点的组成不尽相同，但是从结构上来说一般都包含以下 4 个部分：数据采集、数据处理、数据传输和电源。感知信号的形式通常决定了传感器的类型。而现有的传感器结点的处理器通常包括嵌入式 CPU，如 ARM 公司的 ARM 系列、Motorola 的 68HC16 和 Intel 公司的 8086 等。数据传输单元主要由低功耗、短距离的无线模块组成，如 RFM 公司的 TR1000 等。另外运行于传感器网络上的微型化的操作系统主要负责复杂任务的系统调度与管理，比较常见的有 UC Berkeley 开发的 TinyOS 以及 μCOS-Ⅱ嵌入式 Linux。

如图 3.1 所示是一个典型的传感器体系结构图，传感器模块负责数据的感知和产生

及数模转化，信息处理模块负责进行信号处理，最后经由无线通信模块发射出去。

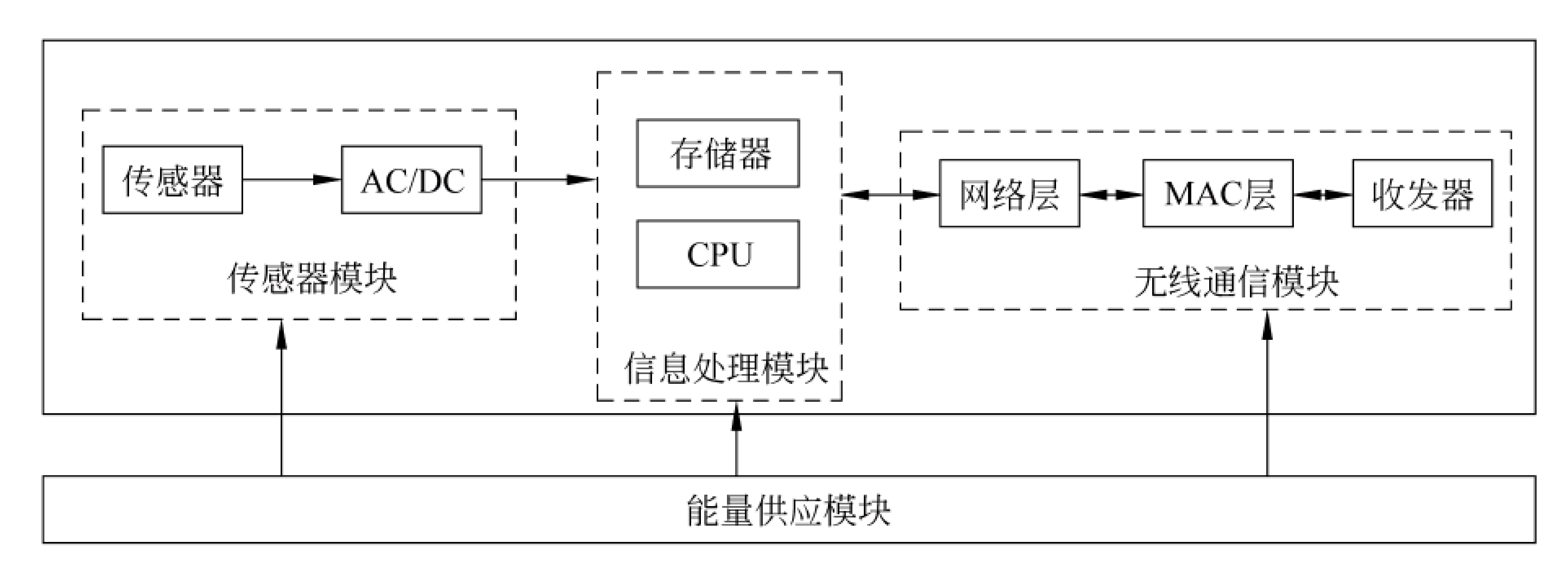


图 3.1　传感器结点体系结构图

传感器网络结点的一些技术参数包括如下几项。

(1) 电池能量。传感器的能量一般由电池提供。一次性电池原则上可工作几年时间。

(2) 传输范围。由于传感器结点能量有限，结点的传输范围只能被限制在一个很小的范围之内(通常是 100 米以内，一般为 1～10 米)，否则会造成传感器的能量枯竭。一些技术(比如数据聚集传输技术)通过先将数据进行聚集，然后传输聚集的结果(而不是每个数据)来减少能量的消耗，帮助减少传感器结点的传输能耗。

(3) 网络带宽。传感器网络的带宽通常只有几十千位每秒。如使用蓝牙协议时小于 723Kbps，使用 802.15.4ZigBee 协议时为 250Kbps。

(4) 内存大小。传感器结点的内存大小一般在 6～8Kb，而且一般的空间被传感器网络的操作系统所占据，例如 TinyOS。内存大小通常会影响密钥管理方案的可行性，即密钥管理方案必须能够有效地利用剩余的存储空间，完成密钥的存储，缓存消息等。

(5) 预先部署的内容。通常，传感器网络具有随机性和动态性，因为不可能获取应用环境的所有情况。预先在传感器结点上配置的信息通常是密钥类的信息，例如，通过预先在结点中存储一些秘密共享密钥，使得网络在部署之后能够实现结点间的安全通信。

2. 典型研究对象

加州大学伯克利分校发起的 smart dust 项目开发了多种传感器结点，如 WeC、Mica、Mica2、MicaZ 等。目前普遍采用的是 2004 年开发的 Telos 结点，采用 16 位 4MHz TI 公司的 MSP430 处理器，正常工作状态下功耗 3mW，该处理器芯片具有 5 种低功耗模式，一般睡眠模式下功耗仅为 225μW，深度睡眠模式下功耗仅为 7.8μW。内存 10KB，闪存 48KB。采用的通信芯片是 Chipcon 公司的 CC2420 通信芯片，工作在 2.4GHz 频道上，符合 IEEE 802.15.4 协议规范，数据传输率达到 250Kbps。

3. 无线传感器网络的网络结构

无线传感器网络在不同的应用场景中的网络拓扑结构可能不同。比较典型的应用方式是：无线传感器结点被任意地散落在监测区域，然后结点间以自组织的形式构建网络，对感知参数进行监测并生成感知数据，最后通过短距离无线通信(如 ZigBee)经过多次转发将数据传送到网关(Sink 结点或者汇聚结点)，网关通过远距离无线通信网络(如

GPRS)将数据发到控制中心。也有传感器结点直接将感知的数据发给控制中心的,这便是一种典型的M2M通信场景。一般而言,无线传感器网络的结构可以分为分布式网络结构和集中式网络结构两种。

1) 分布式无线传感器网络

分布式无线传感器网络没有固定的网络结构,网络拓扑结构在部署前也无法确定。传感器结点通常随机部署在目标区域中。一旦结点被部署,它们就开始在自己的通信范围内,寻找邻居结点,建立数据传输路径。如图3.2所示为分布式网络结构的示意图。

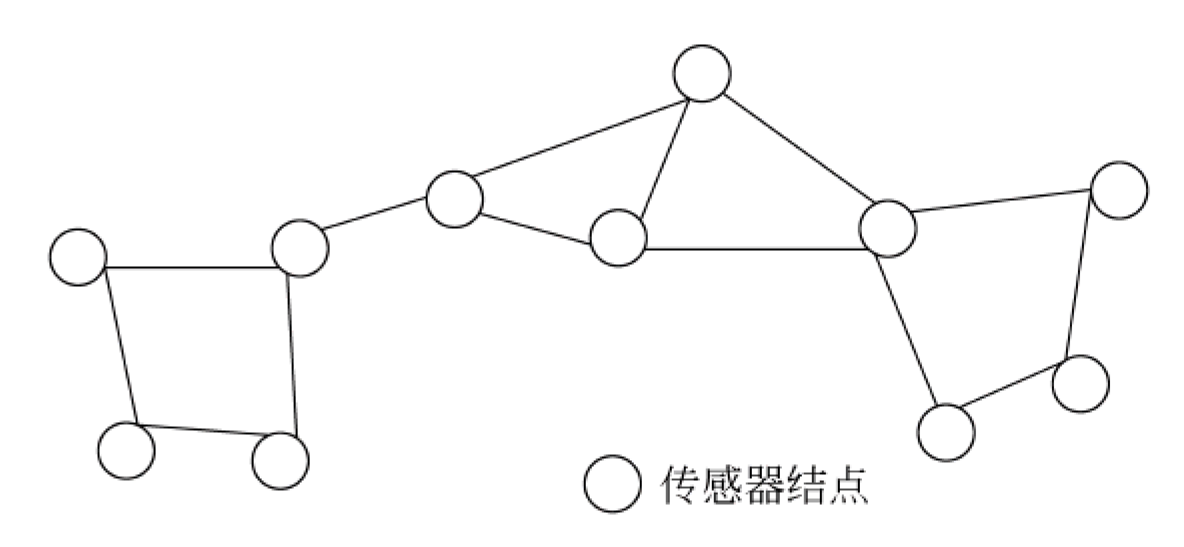


图3.2 分布式网络结构的示意图

2) 集中式无线传感器网络

在集中式无线传感器网络中,依据结点能力的不同可以分为基站、簇头(Cluster Head)结点和普通结点。基站是一个控制中心,通常认为它具有很高的计算和存储能力,可以实施多种控制命令。基站的功能包括以下几种:典型的网络应用中的网关、具有强大的数据存储/处理能力、用户的访问接口。基站通常被认为是抗攻击、可信赖的,因而基站可成为网络中的密钥分发中心。结点通常部署在与基站一跳或多跳的范围内,多跳结点形成一个簇结构(簇结构即包含一个簇头结点和多个普通结点或孩子结点的树状结构)。基站具有很强的传输能力,通常可以与任意一个网络内的结点通信,而结点的通信能力则取决于结点自身的能量水平和位置。依据通信方式的不同,网络内的数据流可以分为点对点通信、组播通信、基站到结点的广播通信。如图3.3所示为集中式网络结构的简图。

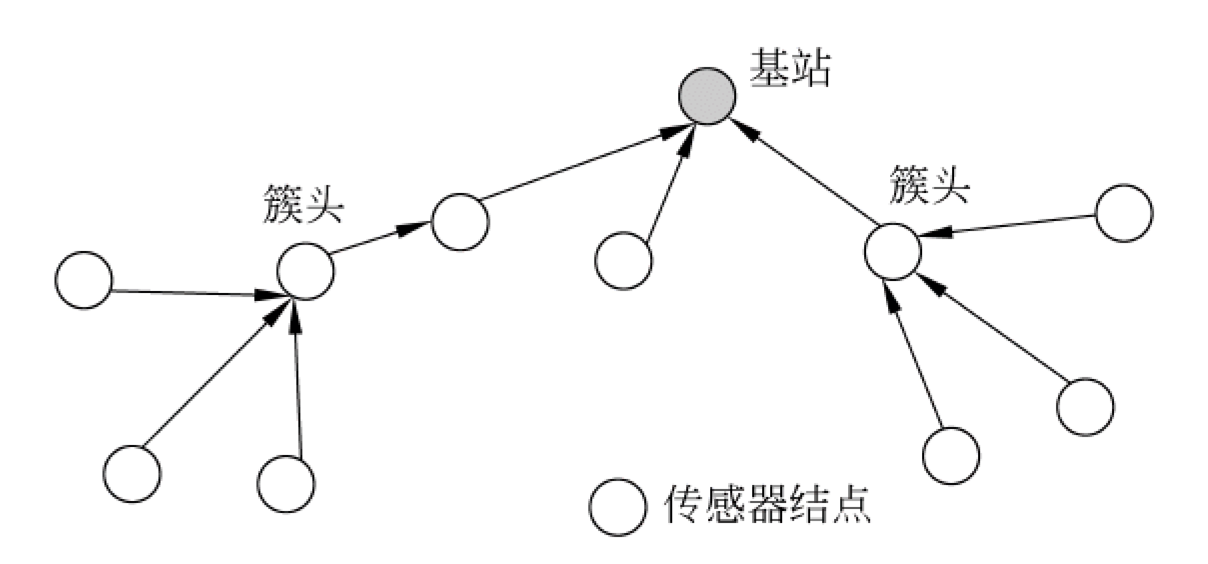


图3.3 集中式网络结构的简图

无线传感器网络的特点如下,在设计安全方案时需要考虑到这些特点:

(1) 网络结点数量众多,结点密度大(即单位面积内的结点数量较多)。

(2) 网络拓扑结构不稳定,拓扑结构随时会发生变化。

(3) 传感器结点受到应用环境和结点成本的限制，计算和通信能力有限。

(4) 能量受限：无线传感器网络由于部署在特定环境中，通常没有持续的外接电能供应，多以电池作为能量源。

3.1.2 无线传感器网络的安全需求分析

通常无线传感器网络会被部署在不易控制、无人看守、边远或易于遭到恶劣环境破坏或者恶意破坏和攻击的环境当中，因而无线传感器网络的安全问题成为研究的热点。由于传感器结点本身计算能力和能量受限的特点，寻找轻量级(计算量小、能耗低)的适合于无线传感器网络特点的安全手段是研究所面临的主要挑战。

1. 安全需求

(1) 通信与储存数据的机密性。无线传感器网络通信不应当向攻击者泄漏任何敏感的信息。在许多应用中，结点之间传递的是高度敏感的数据或者控制信息。结点保存的感知数据、秘密密钥及其他传感器网络中的机密信息(如传感器的身份标识等)，必须只有授权的用户才能访问。同时，因密钥泄漏造成的影响应当尽可能控制在一个小的范围内，从而使得一个密钥的泄漏不至于影响整个网络的安全。解决通信机密性主要依靠使用通信双方共享的会话密钥来加密待传递的消息，解决存储机密性主要依靠加密数据的访问控制。

(2) 消息认证和访问结点认证。结点身份认证在无线传感器网络的许多应用中是非常重要的。例如攻击者极易向网络注入信息，接收者只有通过身份认证才能确信消息是从正确的结点发送过来。数字签名通常不适用于通信能力、计算速度和存储空间都相当有限的传感器结点。传感器网络通常使用基于对称密码学的认证方法，即判断对方是否拥有共享的对称密钥来进行身份的认证。

(3) 通信数据和存储数据的完整性。资源有限的传感器无法支持高计算量的数字签名算法，通常使用对称密钥体制的消息鉴别码来进行数据完整性检验。

(4) 新鲜性。在无线传感器网络中，基站和簇头需要处理很多结点发送过来的采集信息，为防止攻击者进行任何形式的重放攻击(将过去窃听的消息重复发送给接收者，耗费其资源使其不能提供正常服务)，必须保证每条消息是新鲜的。由于密钥可能需要进行更新，因而新鲜性还体现在密钥建立过程中，即通信双方所共享的密钥是最新的。

(5) 可扩展性(Scalability)。这是无线传感器网络的特色之一，由于传感器结点数量大、分布范围广，环境条件、恶意攻击或任务的变化可能会影响传感器网络的配置。同时，结点的经常加入、物理破坏或电量耗尽等也会使得网络的拓扑结构不断发生变化。无线传感器网络的可扩展性表现在传感器结点的数量、网络覆盖区域、生命周期、时间延迟等方面的可扩展程度。因此，给定无线传感器网络的可扩展性级别，例如结点的数量级，安全解决方案必须提供支持该可扩展性级别的安全机制和算法，来使传感器网络保持良好的工作状态。

(6) 可用性(Availability)。无线传感器网络的安全解决方案所提供的各种服务能被授权用户使用，并能有效防止非法攻击者企图中断传感器网络服务的恶意攻击。一个合理的安全方案应当具有节能高效的特点，各种安全协议和算法的设计不应当太复杂，并尽可能地避开公钥密码运算(如公钥加密/解密或者数字签名和签名验证)，计算开销、存储

容量和通信能力、能量消耗的最小化，最终延长网络的生命周期。

(7) 健壮性(Robustness)。无线传感器网络一般配置在恶劣环境或无人区域，环境条件、现实威胁和当前任务具有很大的不确定性。这要求传感器结点能够灵活地加入或去除、传感器网络之间能够进行合并或拆分，因而安全解决方案应当具有鲁棒性和自适应性，能够随着应用背景的变化而灵活拓展，安全解决方案尽可能满足所有可能的应用环境和条件。此外，当某个或某些结点被攻击者控制后，安全解决方案应当限制其安全影响范围，保证整个网络不会因此而失效。

(8) 自组织性(Self-Organization)。由于无线传感器网络是由一组传感器以自组织的(Ad Hoc)方式构成的无线网络，这就决定了相应的安全解决方案也应当是自组织的，即在无线传感器网络配置之前通常无法假定结点的任何位置信息和网络的拓扑结构，也无法确定某个结点的邻近结点集。

2. 安全方案设计时的考虑因素

由于无线传感器网络本身的特点，其安全目标的实现与一般网络不同，在研究和移植各种安全技术时，必须进一步考虑以下约束：

(1) 能量限制。结点在部署后很难替换和充电，所以低能耗是设计安全算法时首要考虑的因素。能耗特点包括：通信芯片耗能占整个传感器结点能耗的比重最大，如常用的 TelosB 结点上，CPU 在正常状态电流只有 500μA，而通信芯片在发送和接收数据时的电流近 200mA。另外，低功耗的通信芯片在发送状态和接收状态消耗的能量差别不大。因而，安全方案应该尽量减少通信(如协议交互)的次数。

(2) 有限的存储、运行空间和计算能力。目前微处理器一般配有 4～10KB 内存，48～128KB 的闪存。

(3) 结点的物理安全无法保证。在进行安全设计时必须考虑被敌手所控制的结点(也称为被俘结点、妥协结点)的检测、撤除问题，来自内部被俘结点发起的攻击，同时还要将被俘结点导致的安全隐患扩散限制在最小范围内。

(4) 结点布置的随机性。结点往往是被随机地投放到目标区域的，结点之间的位置关系一般在布置前是不可预知的。

(5) 通信的不可靠性。无线通信信道的不稳定、结点并发通信的冲突和多跳路由的较大延迟使得设计安全算法时必须考虑容错问题，合理地协调结点通信，并尽可能减少对时间同步的要求。

另外，无线传感器网络的应用十分广泛，而不同的应用场景对安全的需求往往是不同的，应该根据实际的应用来分析具体的安全需求。

3.2 无线传感器网络的安全攻击与防御

3.2.1 常见网络攻击方法

由于传感器网络采用无线通信，开放的数据链路是不安全的，攻击者可以窃听通信的内容，实施干扰。而且传感器结点通常工作在无人区域，缺乏物理保护，容易损坏，且攻击

者可以获取结点，读取存储内容甚至写入恶意代码。攻击通常与使用的数据链路层协议（如 IEEE 802.15.4）、网络层协议（如路由协议、传输层协议）有关。本节首先对各种攻击简单进行分类，然后按网络体系各层归纳各种攻击方法。

（1）阻塞（Jamming）攻击：一种针对无线通信的 DoS 攻击。攻击方法是干扰正常结点通信所使用的无线电波频率，达到干扰正常通信的目的。攻击者只需要在结点数为 N 的网络中随机布置 K（K＜＜N）个攻击结点，使它们的干扰范围覆盖全网，就可以使整个网络瘫痪。

（2）耗尽（Exhaustion）攻击：恶意结点侦听附近结点的通信，当一帧快发送完时，恶意结点发送干扰信号。传统的 MAC 层协议中的控制算法往往会重传该帧，反复重传造成被干扰结点电源很快被耗尽。自杀式的攻击结点甚至一直对被攻击结点发送请求（Request）信号，使得对方必须回答，这样两个结点都耗尽电源。这一攻击的原理可能与具体 MAC 层协议（如 IEEE 802.15.4 协议）有关。

（3）非公平竞争攻击。由于无线信道是单一访问的共享信道，采取竞争方式进行信道的分配，该攻击是指在网络中的某些恶意结点总是占用链路信道，采用一些设置，如较短的等待时间进行重传重试、预留较长的信道占用时间等，企图不公平地占用信道。这一攻击的原理与 MAC 层协议有关。

（4）汇聚结点（Homing）攻击：传感器网络中有些结点执行路由转发功能，Homing 攻击针对这一类结点。攻击者只需要监听网络通信，就可以知道簇头的位置，然后对其发动攻击。簇头瘫痪后，在一段时间内整个簇都不能工作。它也属于 DoS 攻击的一种。

（5）怠慢和贪婪（Neglect and Greed）攻击：其含义是少转发、不转发或多转发收到的数据包。攻击者处于路由转发路径上，但是随机地对收到的数据包不予转发处理。如果向消息源发送收包确认，但是把数据包丢弃不予转发，该攻击称为怠慢（Neglect）。如果被攻击者改装的结点对自己产生的数据包设定很高的优先级，使得这些恶意信息在网络中被优先转发，该攻击称为贪婪（Greed）。

（6）方向误导（Misdirection）攻击：这里的方向是指数据包转发的方向。如果被敌人所控制的路由结点将收到的数据包发给错误的目标，则数据源结点受到攻击；如果将所有数据包都转发给同一个正常结点，则该结点很快因接收包而耗尽电源。方向误导攻击的一个变种是 Smurf 攻击。

（7）黑洞（Black Holes）攻击：又称为排水洞（Sinkholes）攻击。攻击者（用 A 表示）声称自己具有一条高质量的路由到基站，比如广播“我到基站的距离为零”。如果 A 能发送到很远的无线通信距离，则收到该信息的大量结点会向 A 发送数据。大量数据到达 A 的邻居结点，它们都要给 A 发送数据，造成信道的竞争。由于竞争，邻居结点的电源很快被耗尽，这一区域就成了黑洞，通信无法传递过去。对于收到的数据，A 可能不予处理。黑洞攻击破坏性很强，基于距离向量（Distance Vector）的路由算法容易受到黑洞攻击，因为这些路由算法将距离较短的路径作为优先传递数据包的路径。

（8）虫洞（Wormholes）攻击：通常由两个移动主机攻击者合作进行。一个主机 A 在网络的一边收到一条消息，比如基站的查询请求，通过低延迟链路传给距离很远的另一个主机 B，B 就可以直接广播出去，这样，收到 B 广播的结点就会把传感的数据发给 B，因为

收到B广播的结点认为这是一条到达A的捷径。

(9) Hello泛洪(Hello Flood)攻击：在许多协议中，结点通过发送一条Hello消息表明自己的身份，而收到该消息的结点认为发送者是自己的邻居(因为数据包可以到达)。但移动主机攻击者可以将Hello消息传播得很远，远处的正常结点收到消息之后于是把攻击者当成自己的邻居。这些结点会与“邻居”(移动主机攻击者)通信，导致网络流量的混乱。传感器网络中的几个路由协议，如LEACH和TEEN，易受这类攻击，特别是当Hello包中含有路由信息或定位信息。

(10) 女巫(Sybil)攻击：是指一个结点冒充多个结点，它可以声称自己具有多个身份，甚至随意产生多个假身份，利用这些身份非法获取信息并实施攻击。Sybil攻击能破坏传感器网络的路由算法，还能降低数据汇聚算法的有效性。

(11) 破坏同步(Desynchronization)攻击：在两个结点正常通信时，攻击者监听并向双方发送带有错误序列号的包，使得双方误以为发生了丢失而要求对方重传。攻击者使正常通信双方不停地重传消息，从而耗尽电源。

(12) 泛洪攻击(Flooding)：指攻击者不断地要求与邻居结点建立新的连接，从而耗尽邻居结点用来建立连接的资源，使得其他合法的对邻居结点的请求不得不被忽略。

(13) 应用层攻击：如感知数据的窃听、篡改、重放、伪造等。结点不合作行为。对应用层功能如结点定位、结点数据收集和融合等的攻击，使得这些功能出现错误。

表3.1所示给出了无线传感器网络中网络攻击分类的小结。

表3.1 无线传感器网络中网络攻击的分类

分类标准	分　类	说　明
攻击者身份	结点型攻击	攻击者与传感器结点的计算和通信能力相当
	移动主机型攻击	攻击者与移动电脑同级别，危害范围广
攻击来源	外部攻击	攻击者是敌方放置的，可以是结点或移动电脑
	内部攻击	网络中的结点被攻击者所控制，从网络内部发起攻击
攻击发生的协议层次	物理层攻击	阻塞攻击
	数据链路层攻击	耗尽攻击、非公平竞争攻击
	网络层攻击	汇聚结点攻击、怠慢和贪婪攻击、方向误导攻击、黑洞攻击、虫洞攻击、Hello泛洪、女巫攻击
	传输层攻击	破坏同步攻击、泛洪攻击
	应用层攻击	如感知数据的窃听、篡改、重放、伪造等，结点不合作

3.2.2 常用防御机制

对于物理层的攻击(如阻塞(Jamming)攻击)使用扩频通信可以有效地防止。另一对策是，攻击结点附近的结点觉察到Jamming之后进入睡眠状态，保持低能耗。然后定期检查Jamming是否已经消失，如果消失则进入活动状态，向网络通报Jamming的发生。

对于传输层的攻击(如 Flooding),一种对策是使用客户谜题(Client Puzzle),即如果客户要和服务器建立一个连接,必须首先证明自己已经为连接分配了一定的资源,然后服务器才为连接分配资源,这样就增大了攻击者发起攻击的代价。这一防御机制对于攻击者同样是传感器结点时很有效,但是合法结点在请求建立连接时也增大了开销。

对于怠慢和贪婪攻击,可用身份认证机制来确认路由结点的合法性;或者使用多路径路由来传输数据包,使得数据包在某条路径被丢弃后,数据包仍可以被传送到目的结点。

抵抗黑洞攻击可采用基于地理位置的路由协议。因为拓扑结构建立在局部信息和通信上,通信通过接收结点的实际位置自然地寻址,所以在别的位置成为黑洞就变得很困难了。

对付女巫攻击有两种探测方法,一种是资源探测法,即检测每个结点是否都具有应该具备的硬件资源。Sybil 结点不具有任何硬件资源,所以容易被检测出来。但是当攻击者的计算和存储能力都比正常传感器结点大得多时,则攻击者可以利用丰富的资源伪装成多个 Sybil 结点。另一种是无线电资源探测法,通过判断某个结点是否有某种无线电发射装置来判断是否为 Sybil 结点,但这种无线电探测非常耗电。

对于更多的攻击,通常采用加密和认证机制提供解决方案。例如对于分簇结点的数据层层聚集,可使用同态加密、秘密共享的方法。对于结点定位安全,可采取门限密码学,以及容错计算的方法等。然而在无线传感器网络中,传感器结点的计算资源非常有限,通常公钥加密和签名算法因计算量太大而不适用,所以对称密钥加密方案研究得较多,而为了应用对称密钥加密方法,首先需要解决加密密钥的管理问题,这将在下一节介绍。表 3.2 给出了对攻击防御方法的小结。

表 3.2　无线传感器网络攻击防御方法

网络层次	攻击方法	防御方法
物理层	阻塞攻击	扩频、优先级消息、区域映射、模式转换
	物理破坏	破坏感知、结点伪装和隐藏
数据链路层	耗尽攻击	设置竞争门限
	非公平竞争	使用短帧策略和非优先级策略
网络层	丢弃和贪婪攻击	冗余路径、探测机制
	汇聚结点攻击	加密和逐跳(hop-to-hop)认证机制
	方向误导攻击	出口过滤、认证、监测机制
	黑洞攻击	认证、监测、冗余机制
传输层	破坏同步攻击	认证
	泛洪攻击	客户端谜题
应用层	感知数据的窃听、篡改、重放、伪造	加密、消息鉴别、认证、安全路由、安全数据聚集、安全数据融合、安全定位、安全时间同步
	结点不合作	信任管理,入侵检测

3.3 无线传感器网络的密钥管理

无线传感器网络安全中有诸多问题，如安全路由，安全定位，安全数据聚集等，篇幅所限这里只能对一个典型问题加以展开介绍。密钥管理问题是无线传感器网络需要首先解决的安全问题，因为密钥的建立与分发是保密通信的前提。同时，由于传感器结点数量庞大、随机布置的(具有随机网络拓扑结构)，且结点具有资源受限(计算、存储和能量有限)，结点可能因断电、被损坏、被捕获而失效或泄漏密钥，于是密钥管理问题变得更加棘手。因而，密钥管理的可扩展性、自组织性、鲁棒性等要求较高，成为无线传感器网络中一个独具特色的研究问题。

3.3.1 密钥管理的分类与评价指标

传感器结点间共享的秘密密钥是消息加密、消息完整性保护和传感器结点认证的主要依据，因此，如何产生、分发、建立、更新、撤销这些密钥是一个首先需要解决的安全问题。

密钥管理协议分为预先配置密钥协议、有仲裁的密钥协议、分组分簇密钥协议等(这些分类之间可能会有重叠)。预先配置密钥协议即在传感器结点在部署的时候预先分配和安装将来要使用的密钥。这种方法简单，但是在动态无线传感器网络中增加或移除结点的时候，就不灵活。在有仲裁的密钥协议中，存在密钥分配中心(Key Distribution Center，KDC)或者可信第三方(Trusted Third Party，TTP)负责建立密钥，KDC 或 TTP 可以是一个结点或者分散在一组可信任的结点中。分组分簇密钥协议中结点被划分成多个簇，每个簇有能力较强(表现在剩余能量上)的一个或者多个簇头，协助密钥分配中心或者基站共同管理整个无线传感器网络。密钥的初始化分发和管理一般由簇头主持，协同簇内结点共同完成。

1. 预先配置密钥

(1) 网络预分配密钥方法。无线传感器网络整个网络共享一个秘密密钥，所有结点在配置前都要装载同样的密匙。这种方法简单，但是若某个结点的密钥被敌人知道，则整个网络中使用的密钥就暴露了，从而整个网络的通信都失去了保密性。

(2) 结点间预分配密钥方法。在这种方法中，网络中的每个结点需要知道与其通信的所有其他结点的 ID 号，在每两个结点间共享一个独立的秘密密钥。如果每个结点都可能与网络中的其他结点通信，并建立一个共享的秘密密钥，假定结点总量为 n 个，则每个结点要存储 $n-1$ 个密钥，整个网络需要的密钥总量为 $n(n-1)/2$ 个。当结点数量达到几千个时，密钥的数量就比较大了。

2. 有仲裁的密钥协议

仲裁协议假设存在建立密钥的可信第三方(TTP)。根据密钥建立的类型，可分为对称密钥分发协议和公钥分发协议。对称密钥分发通常有密钥分发中心(KDC)。对公钥的分发通常比较容易。

密钥建立协议支持组结点的密钥建立，即建议一组结点之间通信需要使用的密钥。

还有一种分等级的密钥确立协议叫做分层逻辑密钥，在具有相同层次的结点之间的建立密钥关系。

除了上述的分类方法以外还有其他一些分类的方法。表 3.3 给出了其他分类及其描述。

表 3.3　密钥管理名称描述

密钥管理方案名称	描　述
基于主密钥的管理方案	网络中只有单一的密钥，进行加密、解密操作
对(Pairwise)密钥方案	把网络内的通信转化为结点间的通信模式，通过结点对之间的安全实现网络的安全
基于公钥的密钥管理方案	基于公钥技术的密钥管理方案，例如椭圆曲线公钥密码技术 TingECC 在传感器网络中的实现
预共享的密钥管理方案	这是目前研究比较成熟的模型，其中的方案主要有预分配机制、q-composite 机制、多路增强机制、随机预分配方案，以及基于位置信息的密钥管理方案等
动态密钥管理方案	提高了网络的适应能力，更好地支持网络规模的变化
集中式密钥管理方案	主要包括 LEAP 协议[6]、异构传感器网络密钥管理方案等

3. 密钥管理方案的评价指标

评价一种密钥管理技术的好坏，不能仅从能否保障传输数据安全来进行评价，还必须满足如下准则：

(1) 抗攻击性(Resistance)。主要指抗结点妥协的能力。在无线传感器网络中，敌人可能捕获部分结点并复制这些结点来发起新的攻击。针对这种情况，无线传感器网络必须能够抵抗一定数量的结点被捕获而发起的新的攻击。

(2) 密钥可回收性(Revocation)。如果一个结点被敌人控制，对网络产生破坏行为时，密钥管理机制应能采取有效的方式从网络中撤销(Revoke)该结点。撤销机制必须是轻量级的，即不会消耗太多的网络通信资源和结点能量。

(3) 容侵性(Resilience)。如果结点被捕获，密钥管理机制应能够保证其他结点的密钥信息不会被泄漏。即可以容忍网络中被捕获的结点数小于一定的阈值。同时，新结点能够方便地加入网络，参与安全通信。

3.3.2　确定密钥分配方案 Blundo

1. 结点间共享密钥

该模型保证了每个结点之间存在一对共享密钥，结点间会话密钥的建立可以利用该密钥生成。优点是由于要求每个结点必须存储所有其他结点的共享密钥，因而任意两个结点间总可以建立共同的密钥。任何两个结点间的密钥对是独享的，其他结点不知道其密钥信息，任何一个结点被捕获不会泄漏非直接连接的结点的密钥信息。模型简单，实现容易。缺点是扩展性不好，新结点的加入需要更新整个网络中所有的结点所存储的密钥

信息。一旦结点被捕获，敌人可以从结点存储的密钥信息获得该结点与网络所有结点的密钥信息。由于结点需要存储所有其他结点的密钥信息，所以网络规模有限。

2. 结点与基站共享主密钥

网络中的每个结点与基站间共享一对主密钥，每个结点只需要很少的密钥存储空间，基站需要较高的计算和资源开销。优点是对结点的资源和计算能力要求较低，计算复杂度低。密钥建立的成功率高，只要能与基站通信的结点都可以进行安全通信。支持结点的动态更新。缺点是过分依赖基站的能力，基站是单一失效点，即一旦基站被捕获，整个网络即陷入瘫痪。网络的规模取决于基站的通信能力。基站会成为整个网络的通信瓶颈。多跳通信时，结点只负责透明地转发数据包，没有办法对信息报进行任何认证，恶意结点容易利用这一特点，进行 DoS 攻击。

为减少结点间共享主密钥的存储空间，Blundo 提出基于对称二元多项式的方案。

3. Blundo 二元多项式方案

具体而言，设在公开信道上有 n ($n>2$)个用户，每对用户之间要建立一个可进行秘密通信的会话密钥。TA 是一个可信的第三方，一个“平凡的”解决方法是，对于任何一对用户{U,V}，TA 选择一个随机密钥 $K_{UV}=K_{VU}$，并通过“离线”的安全信道传送 U 和 V。但是这种方法每个用户必须存储 $n-1$ 个密钥，且 TA 需要安全地传送 $C_n^2=n(n-1)/2$ 个密钥。当网络用户数量较多时，这一代价是很高的，因而不是一个实用的解决方案。

Blundo 在 1993 年利用二元 t 次多项式提出了对(pairwise)密钥分发模型，其中多项式 $f(x,y)=\sum_{i,j=0}^{t}a_{ij}x^iy^j$ 具有对称性 $f(x,y)=f(y,x)$。结点部署前，任一结点 m 将其身份 ID_m 作为参数替换多项式其中的一个变量的结果 $f(ID_m, y)$存储。那么在部署后，只要某结点 n 知道结点 m 的 ID，便能建立对密钥 $f(ID_m, ID_n)=f(ID_n, ID_m)$。

Blundo 方案的巧妙之处是利用了关于 x 和 y 的多项式的对称性：对于所有的 x、y，$f(x,y)=f(y,x)$，这一性质可被用来构造共享的密钥。

步骤如下：

(1) 公开参数选择：TA 选定一个大素数 $p(p\geqslant n)$，每个用户 U 各自选定一个正整数 $r_u\in Z_p^*$，它们各不相同，TA 公开这些 r_u。

(2) TA 随机选定 3 个 $a,b,c\in Z_p^*$，并构造函数 $f(x,y)=(a+b(x+y)+cxy)\bmod p$。

(3) 对每个用户 U，TA 计算多项式 $g_u(x)=f(x,r_u)$，并将 $g_u(x)$通过安全信道发送给 U。容易得到，$g_u(x)=a_u+b_ux$，其中 $a_u=(a+br_u)\bmod p$，$b_u=(b+cr_u)\bmod p$。

如果 U 要与 V 进行秘密通信，那么 U 和 V 分别计算 $K_{UV}=g_u(r_v)\bmod p$，以及 $K_{VU}=g_v(r_u)\bmod p$。

由于 $K_{VU}=g_u(r_v)\bmod p=f(r_u,r_v)\bmod p=g_v(r_u)\bmod p=K_{vu}$，所以 U 与 V 得到一个共享的密钥 $K_{UV}=K_{VU}$。

例 假设有 3 个用户 u、v 和 w，$p=17$，用户的公开信息为 $r_u=12$，$r_v=7$ 及 $r_w=1$。假定 TA 选择 $a=8$、$b=7$ 和 $c=2$，于是多项式 f 为 $f(x,y)=8+7(x+y)+2xy$。

多项式 g 表示为

$$g_u(x)=7+14x$$

$$g_v(x)=6+4x$$

$$g_w(x)=15+9x$$

由此产生的 3 个密钥为：$K_{uv}=3$、$K_{uw}=4$、$K_{vw}=10$。容易验证：

用户 u 计算 $K_{uv}=g_u(r_v)=7+14\times7 \bmod 17=3$

用户 v 计算 $K_{uv}=g_v(r_v)=6+4\times12 \bmod 17=3$

其余密钥的计算留作练习。

*3.3.3　随机密钥分配方案 EG

上节介绍的密钥管理方案为确定性方案，作为对本节内容的引导，便于读者理解密钥管理方案的基本目的。为进一步减少在结点上存储密钥所需要的空间，提出了随机密钥预分发方案。随机密钥预分配方案最早由 Eschenauer 和 Gligor 提出[7]，因此该方案也被称为 EG 方案。该方案的基本思想是首先建立一个比较大的密钥池，任何结点都拥有密钥池中一部分密钥，那么任意两个结点间能够以一定的概率使得双方拥有一对相同的密钥，从而建立安全通道。其思路可简单形象地比喻为：每个结点各自从一堆(pool)密钥中随机取出一串(ring)密钥，结点间将会以一定概率共享一对密钥。即使没有也会从多跳间隔的发送端和接收端间共享路径密钥。

EG 方案的实施分为以下几个阶段：密钥预分发、共享密钥的发现、路径密钥对的建立和密钥的撤销机制。具体描述如下。

(1) 密钥预分发。在一个比较大的密钥空间内，为一个无线传感器网络选择一个密钥池，并为选中的密钥池中的密钥附加一个唯一的 ID。在密钥预分发时，从密钥池中任意选择 m 个密钥部署在每个结点中。这 m 个密钥构成一个结点的密钥环，结点密钥环的大小，根据结点存储能力而定。

(2) 共享密钥的发现。结点密钥预分发完成之后，它们就被部署到预期的位置，比如医院、战场等一些实际的应用场景。部署之后，每一个结点开始利用与周围结点共享的密钥，寻找自己的邻居结点。寻找邻居结点的方式有很多种，最简单的就是：结点通过广播自己的密钥 ID，寻找与自己有共享密钥的邻居结点，如果结点发现其他结点与自己有共同的密钥，则利用该共享密钥与其建立安全通信链路。但这种方法增加了网络传输开销，在能量受限的传感器网络中不可取。

(3) 路径密钥的建立。在随机预分配模型下，只有当两个结点存在共享对密钥时，才可以进行通信。当两个结点间不存在共享密钥时，可通过在结点间建立一个路径密钥，从而建立安全通信链路。例如，假定结点 U 希望与结点 V 通信，但是两个结点间不存在共享密钥。U 首先与它的一个邻居结点 Z 发出信息，表示希望与 V 通信。如果结点 Z 与结点 V 存在共享密钥，则结点 Z 生成一个共享密钥 K_{UV} 分别发送给结点 U 和 V。此时，结点 Z 的作用可以视为是一个中介，或者是密钥分发中心。经过共享密钥的发现和路径密钥对的建立，网络中的结点可以相互之间安全通信。同时，由于共享密钥的建立是在各自拥有的密钥的基础之上，因而网络的安全性得到保障。

(4) 密钥撤销机制。在该模型下，考虑结点被捕获(妥协)的情况。由于每个结点包含有一定数量的密钥信息，因而网络的安全受到一定的威胁。为了应对结点捕获，网络中

的其他结点必须能够删除与被捕获结点间的共享密钥。为了能够检测被捕获的结点，EG方案中设定了控制结点，该结点的功能类似于一些方案中的基站。它具有很高的安全性和可信赖性，能够检测出被捕获的结点。同时，该方案还假设在结点部署前，控制结点与网络中的所有结点具有共享密钥。EG 方案的改进 Q-Composite 可以参阅文献[8]。

3.4 无线传感器网络安全协议 SPINS

无线传感器网络安全协议 SPINS 是无线传感器网络安全框架之一，它有两个模块：SNEP(Security Network Encryption Protocol)和 uTELSA。SNEP 提供的安全保障是：数据机密性、数据鉴别和数据新鲜度保证。不同于其他网络，无线传感器网络往往通过广播方式工作，因此低能耗的广播鉴权(认证)协议是一个十分重要的研究问题。uTELSA 提供了在资源受限情况下的广播认证。

3.4.1 轻量级安全协议 SNEP

SNEP 是一种低开销安全协议，保证了数据机密性、数据鉴别(完整性保护)、数据新鲜度等。SNEP 本身只描述协议的过程，没有规定实际采取的具体算法，具体实现时可以根据实际情况选用不同的算法。

SNEP 中，每个结点都和基站之间共享一对主密钥 K_{master}，其他密钥通过使用主密钥来生成。SNEP 的优点是具有较低的通信开销，每条消息仅仅增加 8 字节；使用了计数器，避免了计数器值的传递；加密机制具有较高的安全性(如语义安全)，能够阻止窃听者从被加密的信息中推导出信息的内容。最后，该协议也提供了数据认证、重放保护以及消息新鲜度的功能。

1. SNEP 中的数据机密性

在使用 CTR(Counter)模式时，通信双方共享一个计数器，计数器的值作为每次通信加密的初始化向量，由于每次通信时的计数器的值都不同，于是即使是相同的明文被加密也会导致不同的密文。通信双方共享计数器，在每个分组之后增加，于是避免了因传递计数器值而导致的能量消耗。加密的数据格式是：

$$E=\{D\}_{(Ke,C)}$$

其中，Ke 是加密密钥，C 是计数器的值，作为 CTR 模式的初始化向量。

2. SNEP 中的数据完整性

通过消息鉴别码(MAC)完成，有

$$MAC=\{C \parallel E\}_{Kmac}$$

其中，C 是计数器值，E 是密文，K_{mac}是数据完整密钥。这是一种密文鉴别的方法，即直接验证密文的完整性，避免了不必要的解密运算。

K_e 和 K_{mac}都是从主密钥 K_{master}生成的，生成的方式可以依据实际情况选定，只要在通信双方均实现了该生成算法。例如，可利用 uTELSA 中定义的单向密钥生成函数 F 来生成这两个密钥：

$$K_e = F^{(1)}(K_{master}), \quad K_{mac} = F^{(2)}(K_{master})$$

3. SNEP 中消息的新鲜性

消息的新鲜性可防御重放攻击，SNEP 采用的强新鲜性认证使用了 Nonce 机制，Nonce 是一个使用一次的且无法预测的随机值，通常由伪随机数生成器产生。在结点 A 发送给结点 B 的消息中，包含了 Nonce 值 N_A，在 B 对该消息的应答中，需要包含该值。如下：

$$A \to B: N_A, \{RQST\}_{(Ke,C)}, \{C \parallel \{RQST\}_{(Ke,C)}\}_{Kmac}$$

$$B \to A: \{RPLY\}_{(Ke,C')}, \{N_A \parallel C' \parallel \{RPLY\}_{(Ke,C')}\}_{Kmac}$$

其中，RQST 是请求包，RPLY 是应答包。

4. 结点间的安全通信

SPINS 中每个结点与基站(或者是 Sink 结点、网关等)之间共享一个主密钥，对于结点上传数据到基站的应用，这一方式是可行的，但在有些应用中，结点间或者簇内也需要通信，如果都经过基站转发则效率较低。一个可行的办法就是通过基站建立结点间的临时通信密钥，这里基站起了对称密钥分配中心(KDC)的作用。例如结点 A 和 B 之间需要通信，可采取的方式如下：

$$A \to B: N_A, A$$

$$B \to S: N_A, N_B, A, B, \{N_A \parallel N_B \parallel A \parallel B\}_{K_{BS}}$$

$$S \to A: \{SK_{AB}\}_{K_{AS}}, \{N_A \parallel B \parallel \{SK_{AB}\}_{K_{AS}}\}_{K_{AS}}$$

$$S \to B: \{SK_{AB}\}_{K_{BS}}, \{N_A \parallel B \parallel \{SK_{AB}\}_{K_{BS}}\}_{K_{BS}}$$

其中，K_{AS}是 A 与基站 S 之间的共享主密钥，K_{BS}是 B 与基站 S 之间的共享主密钥，SK_{AB}是结点 A 和 B 之间将要建立的新的临时通信密钥。N_A和 N_B是 Nonce。

3.4.2 广播认证协议 uTELSA

无线传感器网络中，基站通常采用广播方式查询结点，结点收到广播包后，需要对广播包的来源进行认证，若通过认证再进行回复。若采取对称密钥，广播认证和单播认证的区别在于：单播包的认证依赖于收、发结点间共享的一个密钥，而广播包认证需要全网络共享一个公共密钥。这导致安全性较差，即任何一个结点被俘虏将会泄漏整个网络的广播认证密钥。如果采取密钥更新的方法来更新广播认证密钥，需要增加通信开销。传统的广播认证往往依赖于非对称密钥，即利用发送者的数字签名，接收者用公钥进行验证。但是这种方式对于传感器网络而言开销太大，签署签名和验证签名的计算量较大，签名的传递也导致额外的通信负担。针对传感器网络的广播认证问题，Adrian Perrig 等人设计了 uTELSA 协议。该协议使用对称密钥机制实现了一个轻量级的广播认证。

uTELSA 要求基站和结点间保持松散的时间同步，每个结点都知道最大同步误差的上限。为了发送广播认证包，基站计算该包的 MAC，使用的是该时间段的密钥。当一个结点收到该广播认证包时，通常认为验证该包 MAC 的密钥还没有被基站透露。既然只有基站拥有该密钥，所以可认为该包是没有被攻击者改变的。结点存储该数据包在缓存中，等待基站透露验证 MAC 的密钥。基站于是广播验证 MAC 的认证密钥给所有的接

收者，结点接收到该密钥后，便可以验证缓存中的那个广播包的 MAC 的正确性。

MAC 密钥都是密钥链中的一个密钥，密钥链是通过一个单向函数 F 产生的。基站要事先生成这样一个密钥链，方法是：使用单向函数 F 计算 $K_i = F(K_{i+1})$。密钥链中的每个密钥都对应一个时间段，所有在同一时间间隔的广播包都使用同一个密钥进行认证。如图 3.4 所示，在两个时间间隔后，相应的密钥才透露。密钥透露是一个独立的广播数据包。

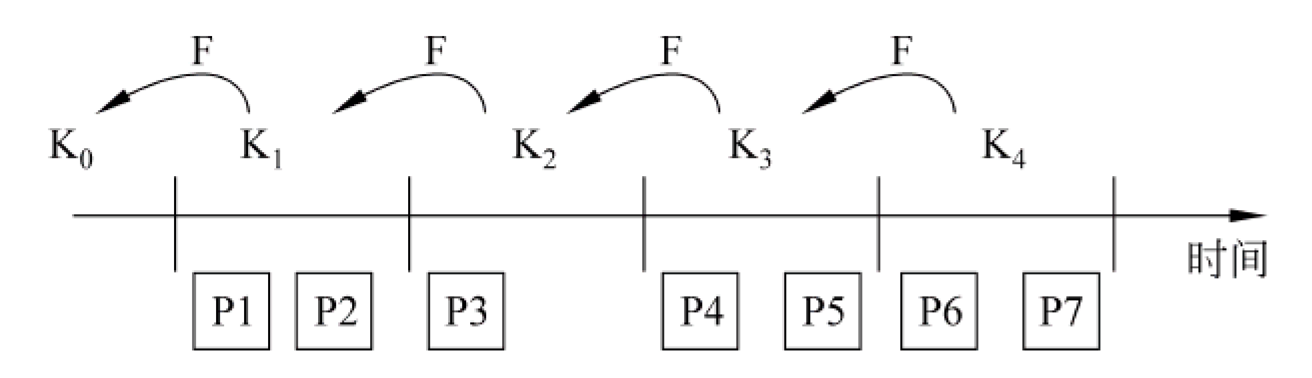


图 3.4　时间松散同步的广播认证

假设接收结点是大体上与基站时间同步的，并知道初始密钥 K_0，数据包 P1、P2 中的 MAC 由密钥 K_1 生成，在时间间隔 1 内发送。

数据包 P3 中的 MAC 由 K_2 生成，在时间间隔 2 发送。此时接收者不能认证任何数据包，因为 K_1 要到时间间隔 3 才透露。

类似地，数据包 P4 和 P5 的 MAC 由 K_3 生成，在时间间隔 3 发送。假设数据包 P4 和 P5 丢失了，同时透露密钥 K_1 的包也丢失了，则接收者仍然不能验证 P1 和 P2 的完整性（因为没有 K_1）。

在时间间隔 4，基站广播了密钥 K_2，结点可通过验证 $K_0 = F(F(K_2))$，并得到 $K_1 = F(K_2)$，这时可用 K_1 来验证 P1 和 P2 的完整性，利用 K_2 来验证 P3 的完整性。

*3.4.3　轻量级公钥密码算法 NTRU

本节简单介绍一个适合在资源受限系统中使用的轻量级公钥加密算法 NTRU（Number Theory Research Unit），它被认为是实现空间最小的公钥加密算法（约 8KB），可用于传感器结点等嵌入式系统中，甚至是 RFID 标签上。NTRU 算法同时也是 IEEE 1363.1 公钥加密算法标准的一部分。NTRU 算法用到了一些抽象代数中的基本概念，如多项式环。

NTRU 公开密钥算法是一种快速公开密钥体制，于 1996 年在密码学顶级会议 Crypto 会议上由美国布朗大学的 Hoffstein、Pipher、Silverman 三位数学家提出。经过几年的迅速发展与完善，该算法的密码学领域中受到了高度的重视并在实际应用（如无线传感器网络中的加密）中取得了很好的效果。现在还有研究人员试图将其用到 RFID 系统的加密中。

NTRU 是一种基于多项式环的密码系统，其加密、解密过程基于环上多项式代数运算和对数 p 和 q 的模约化运算，由正整数 N、p、q 以及 4 个 N−1 次整系数多项式（f，g，r，m）集合来构建。N 一般为一个大素数，p 和 q 在 NTRU 中一般作为模数，这里不需要保证 p 和 q 都是素数，但是必须保证 $\gcd(p,q)=1$，而且 q 比 p 要大得多。$R=Z[X]/(X^N-1)$ 为多项式截断环，其元素 $f(f \in R)$ 为：$f = a_{N-1}x^{N-1} + \cdots + a_1x + a_0$。定义 R 上多项式元

素加运算为普通多项式之间的加运算,用符合+表示,R 上多项式元素乘法运算为普通多项式的乘法运算,当然乘积结果要进行模多项式 x^N-1 的运算,即 2 个多项式的卷积运算,称为星乘,用⊗表示。R 上多项式元素模 q 运算就是把多项式的系数作模 q 处理,用 mod q 表示。

1. NTRU 密码体制描述

(1) 密钥生成。随机选择两个 N-1 次多项式 f 和 g 来生成密钥。利用扩展的 Euclidean 算法对 f 求逆。如果不能求出 f 的逆元,就重新选取多项式 f。用 F_p,F_q 表示 f 对 p 和 q 的乘逆。即:$F_q \otimes f \equiv 1 \bmod q$,$F_p \otimes f \equiv 1 \bmod p$。

计算:$h \equiv F_q \otimes g \bmod q$

最后得:公钥为(N,p,q,h),私钥为(f, F_p)。

这里 F_p 可以从 f 容易地计算得到,但仍然作为私钥存储,这是因为在解密时需要使用这个多项式,而 F_q 和 q 就不需要存储了。

(2) 加密算法。首先把消息表示成次数小于 N 且系数的绝对值至多为(p-1)/2 的多项式 m,然后,随机选择多项式 r∈L,并计算:$c \equiv (pr \otimes h + m) \bmod q$。密文是多项式 c。

(3) 解密算法。收到密文 c 后,可以使用私钥(f, F_p)对密文 c 进行解密。依次计算:

$$a \equiv (f \otimes c) \bmod q, a \in (-q/2, q/2)$$

$$b \equiv a \bmod p$$

$$m \equiv F_p \otimes b \bmod p$$

正确性证明:由于

$$\begin{aligned} a &\equiv f \otimes c \bmod q \equiv (f \otimes (pr \otimes h + m) \bmod q) \bmod q \\ &\equiv (f \otimes pr \otimes h + f \otimes m) \bmod q \\ &\equiv (f \otimes pr \otimes F_q \otimes g + f \otimes m) \bmod q \\ &\equiv (pr \otimes g + f \otimes m) \bmod q \end{aligned}$$

又因为 a 的系数在区间(-q/2,q/2),所以 $pr \otimes g + f \otimes m$ 的系数在区间(-q/2,q/2),故 $pr \otimes g + f \otimes m$ 模 q 后结果不变。因此

$$\begin{aligned} F_p \otimes b \bmod p &\equiv (F_p \otimes a \bmod p) \bmod p \equiv F_p \otimes (pr \otimes g + f \otimes m) \bmod p \\ &\equiv (F_p \otimes pr \otimes g + F_p \otimes f \otimes m) \bmod p \equiv m \bmod p \end{aligned}$$

从而解密成功。

非正式地说,该加密算法的设计思路是,利用随机多项式 r 生成一个“密钥多项式 h”,利用这个密钥多项式进行加密得到密文多项式。解密时利用多项式取模,约去随机多项式 r,利用多项式的逆,解出明文多项式。可见,同一个明文在不同的加密中会产生不同的密文。

例　设(N,p,q)=(5,3,16),以及 $f=X^4+X-1$ 和 $g=X^3-X$,求公钥私钥对以及描述加密解密过程。

由于$(X^4+X-1) \otimes (X^3+X^2-1) \equiv 1 \bmod 3$,故有 $F_q=X^3+X^2-1$,同理可求得 $F_q=X^3+X^2-1$。又由于 $h \equiv F_p \otimes g \bmod 16 \equiv -X^4-2X^3+2X^2+1$,所以公钥为(N,p,q,h)=$(5,3,16,-X^4-2X^3+2X^2+1)$;私钥为$(f,F_p)=(X^4+X-1, X^3+X^2-1)$。

加密过程：首先将消息 m 表示成多项式 $m=X^2-X+1$，然后选取多项式 $r=X-1$，则密文为：$c\equiv 3r\otimes h+m\equiv -3X^4+6X^3+7X^2-4X-5 \bmod 16$。

解密过程：首先计算 $a\equiv f\otimes c=4X^4-2X^3-5X^2+6X-2 \bmod 16$，计算 $F_p\otimes a\equiv X^2-X+1 \bmod 3$，这样就恢复了消息 m。

2. 注解

解密过程有时候可能无法恢复出正确的明文，因为：

在解密过程

$$a'\equiv(f\otimes c) \bmod q\equiv f\otimes(pr\otimes h+m) \bmod q\equiv(pr\otimes g+f\otimes m) \bmod q$$

中，如果多项式 $pr\otimes g+f\otimes m$ 的系数不在区间 $(-q/2,q/2)$，则

$$f\otimes(pr\otimes h+m) \bmod q\neq pr\otimes g+f\otimes m$$

设 $f\otimes(pr\otimes h+m)=pr\otimes g+f\otimes m+qu$，u 为多项式，并且 u 的系数不全为 0，计算：

$$\begin{aligned} e'&\equiv F_p\otimes a' \bmod p\equiv F_p\otimes(pr\otimes g+f\otimes m+qu) \bmod p \\ &\equiv F_p\otimes pr\otimes g+F_q\otimes f\otimes m+F_p\otimes qu \bmod p \end{aligned}$$

由于 p 和 q 互素，所以 $e'\equiv m+F_p\otimes qu \bmod p\neq m$，所以解密失败。

通过选择恰当的参数 N、p、q 就能够避免以上错误，例如取 $(N,p,q)=(107,3,64)$ 和 $(N,p,q)=(503,3,256)$，实验表明解密错误的概率小于 5×10^{-5}，这就是通常能正确解密的原因。

3. 安全性

NTRU 算法的安全性是基于数论中在一个具有非常大的维数的格(Lattice)中寻找最短向量(Shortest Vector Problem，SVP)是困难的。所谓格是指在整数集上的一个基向量组的所有线性组合的集合。目前解决这个问题的最有效方法是 1982 年提出的 LLL(Lenstra-Lenstra-Lovasz)算法，但该算法也只能解决维度在 300 以内的。只要恰当地选择 NTRU 的参数，其安全性与 RSA，ECC 等加密算法是一样安全的。表 3.4 给出了 NTRU、RSA 以及 ECC 安全强度的比较。

表 3.4　NTRU、RSA、ECC 的安全性比较(密钥长度的比较)

NTRU/位	RSA/位	ECC/位
167	512	113
251	1024	163
347	2048	224
503	4096	307

4. 效率

由于 NTRU 只包括小整数的加、乘、模运算，在相同安全级别的前提下，NTRU 算法的速度要比其他公开密钥体制如 RSA 和 ECC 要快的算法快得多，产生密钥的速度也很快，密钥的位数也较小，存储空间也较少。例如，对于长度为 n 的加密明文(解密密文)，NTRU 需要的运算量为 $O(n^2)$，而 RSA 为 $O(n^3)$。因此，NTRU 算法可降低对带宽、处理器、存储器的性能要求，这使得其在智能卡、无线通信等应用中有实体认证与数字签名

的需求时，NTRU 公钥密码算法是目前一个很好的选择。NTRU 已被接受为 IEEE 1363 标准。表 3.5 给出一些效率的比较。

表 3.5　NTRU 与 RSA 以及 ECC 的运算次数比较

公钥体制	基本运算	需要的运算次数	
		加密	解密
NTRU	卷积	1	2
RSA	模乘	17	≈1000
ECC	椭圆曲线上有理点标量乘	≈160	≈160

注意，NTRU 与 ECC 的基本运算耗时大致相同，而 RSA 的基本运算耗时则相对少一些。

研究与思考

[1]　RFID 和 WSN 融合后的安全问题与对策。
[2]　大规模移动且 IP-enabled WSN 中的密钥管理问题。
[3]　实现 NTRU 算法，并移植到传感器结点上进行性能评价。
[4]　思考具有主动行为能力的 WSN 或 CPS 系统中的密钥管理问题。

进一步阅读建议

WSN 安全方面的文献已经有很多，除了密钥管理外，安全路由、安全定位、安全数据聚集、访问控制、隐私保护等方面都有大量研究成果。面向物联网应用的 WSN 通常具有大量的结点，且可以通过 Internet 访问。

[1]　A. Perrig，J. Stankovic，D. Wagner，Security in Wireless Sensor Networks [J]，Communication of the ACM，47(6)：53-57. 2002.

本章参考文献

[1]　陈娟，张宏莉. 无线传感器网络安全研究综述[J]. 哈尔滨工业大学学报，2011(07).
[2]　李平，林亚平，曾玮妮. 传感器网络安全研究[J]. 软件学报，2006(12).
[3]　张楠，无线传感器网络安全技术研究[M]，成都：西南交通大学出版社，2010.
[4]　沈玉龙，裴庆祺，马建峰，庞辽军. 无线传感器网络安全技术概论[M]. 北京：人民邮电出版社，2010.
[5]　杨庚，陈伟，曹晓梅. 无线传感器网络安全[M]. 北京：科学出版社，2010.
[6]　S. Zhu，S. Setia，S. Jajodia，LEAP：Efficient Security Mechanisms for Large-Scale Distributed Sensor Networks[C]，In Proc. of ACM CCS03，2003，62-72.

[7] L. Eschenauer，V. D. Gligor，A Key Management Scheme For Distributed Sensor Networks[C]，In Proc. of ACM CCS02，2002.

[8] H. Chan，A. Perrig，and D. Song. Random Key Predistribution Schemes for Sensor Networks [C]. In Proc. of the IEEE S&P03，2003，197-213.

[9] Wei Ren，Yi Ren，Hui Zhang，Secure，Dependable and Publicly Verifiable Distributed Data Storage in Unattended Wireless Sensor Networks[J]，Springer SCIENCE CHINA Series F，Springer，March 2010，Vol. 53 No. 3：677-692.

[10] Wei Ren，Junge Zhao，Yi Ren，Network Coding based Dependable and Efficient Data Survival in Unattended Wireless Sensor Networks[J]，Journal of Communications，Academy，2009，4(11)：894-901.

第 4 章　物联网终端系统安全

物联网感知层存在大量的终端设备，包括第 2 章介绍的 RFID 标签、读写器等，以及第 3 章介绍的传感器结点。

尤其值得注意的是，智能手机可以说是一种随身携带的“超级”感知和识别设备。智能手机上可以配备的传感器种类繁多：加速度传感器、陀螺仪传感器、温度传感器，地磁传感器、方向传感器、压力传感器、距离传感器、光线亮度传感器等。手机具备 GPS 定位功能，可提供基于位置的服务。手机上的摄（照）像功能也是感知声音、图像、影像能力的体现。加上语音识别和手写字体识别，表现为一种识别能力。如果 RFID 标签附着在手机内部，手机便具有标识（手机使用者）的功能，于是产生了手机门票。手机触摸屏装有 RFID 读写器，于是手机具有标签读取功能（例如苹果公司已经申请了该专利用于下一代 iPhone 产品），于是手机可用于读取标签识别物体的能力。

广义而言，物联网终端通常可分为两种：一种感知识别型终端，以二维码、RFID、传感器为主，实现对“物”的识别或环境状态的感知；另一种就是应用型终端，包括输入/输出控制终端，如计算机、平板电脑、智能手机等终端。感知识别型终端的系统安全中，以嵌入式系统的安全问题为代表；应用型终端的系统安全问题中，以智能手机的安全问题为重中之重。因此，本章重点介绍嵌入式系统安全和智能手机系统安全。

4.1　嵌入式系统安全

4.1.1　嵌入式系统的安全架构

物联网的感知识别型终端系统通常是嵌入式系统。所谓嵌入式系统，是以应用为中心，以计算机技术为基础，并且软/硬件是可定制的，适用于对功能、可靠性、成本、体积、功耗等有严格要求的专用计算机系统。嵌入式系统的发展经历了无操作系统、简单操作系统、实时操作系统和面向 Internet 等 4 个阶段。嵌入式系统的典型结构如图 4.1 所示。

结合嵌入式信息系统的结构，从硬件平台、操作系统和应用系统 3 个方面对嵌入式系统的安全性加以分析。

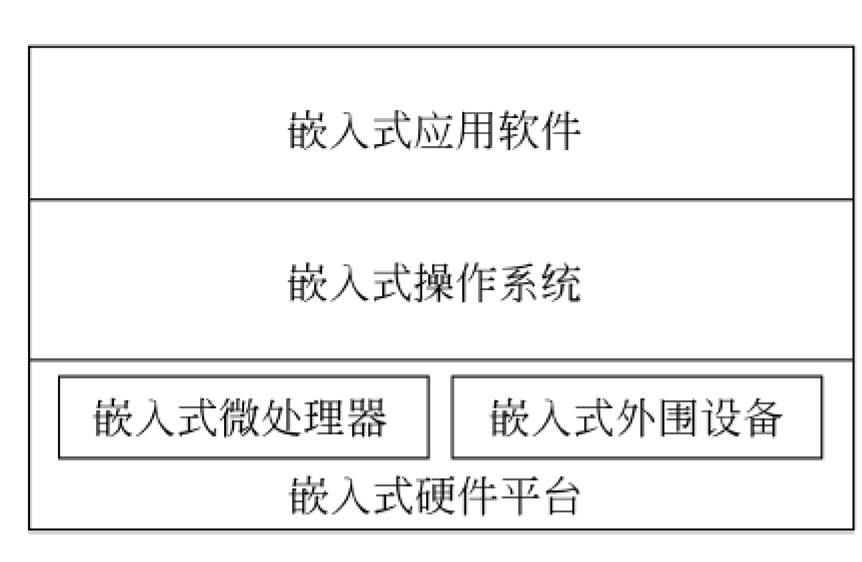


图 4.1　嵌入式系统的典型结构

1. 硬件平台的安全性

为适应不同应用功能的需要，嵌入式系统采取多种多样的体系结构，攻击者可能采取的攻击手段也呈现多样化的特点。区别于 PC 系统，嵌入式信息系统可能遭到的攻击存在于系统体系结

构的各个部分。

(1) 对可能发射各类电磁信号的嵌入式系统,利用其传导或辐射的电磁波,攻击者可能使用灵敏的测试设备进行探测、窃听甚至拆卸,以便提取数据,导致电磁泄漏攻击或者侧信道攻击。而对于嵌入式存储元件或移动存储卡,存储部件内的数据也容易被窃取。

(2) 针对各类嵌入式信息传感器、探测器等低功耗敏感设备,攻击者可能引入极端温度、电压偏移和时钟变化,从而强迫系统在设计参数范围之外工作,表现出异常性能。特殊情况下强电磁干扰或电磁攻击则可能将毫无物理保护的小型嵌入式系统彻底摧毁。

2. 操作系统的安全性

与 PC 不同的是,嵌入式产品采用数十种体系结构和操作系统,著名的嵌入式操作系统包括 Windows CE、VxWorks、pSoS、QNX、PalmOS、OS-9、LynxOS、Linux 等,这些系统的安全等级各不相同,但各类嵌入式操作系统普遍存在因运行的硬件平台计算能力和存储空间有限,于是精简代码而牺牲其安全性的情况。嵌入式操作系统普遍存在的安全隐患如下:

(1) 由于系统代码的精简,对系统的进程控制能力并没有达到一定的安全级别。

(2) 由于嵌入式处理器的计算能力受限,缺少系统的身份认证机制,攻击者可能很容易破解嵌入式操作系统的登录口令。

(3) 大多数嵌入式操作系统文件和用户文件缺乏必要的完整性保护控制。

(4) 嵌入式操作系统缺乏数据的备份和可信恢复机制,系统一旦发生故障便无法恢复。

(5) 各种嵌入式信息终端病毒正在不断出现,并大多通过无线网络注入终端。

3. 应用软件的安全性

应用软件的应用层面的安全问题,如病毒、恶意代码攻击等,应用软件的中间件安全问题,应用软件的系统层面(如网络协议栈)的安全问题,如数据窃听、源地址欺骗、源路由选择欺骗、鉴别攻击、TCP 序列号欺骗、拒绝服务攻击等,也同样存在。

4. 嵌入式系统安全的对策

通常嵌入式系统安全的对策可根据安全对策所在位置分为 4 层,如图 4.2 所示,下面分别加以解释。

安全应用层(应用程序、网络安全协议等)
软件安全架构层(操作系统、虚拟机等)
硬件安全架构层 (处理器、内存、加密处理器等)
安全电路层(电路元件、封装等)

图 4.2 嵌入式系统的分层安全对策

(1) 安全电路层。通过对传统的电路加入安全措施或改进设计,实现对涉及敏感信息的电子器件的保护。一些可以在该层采用的措施主要有:通过降低电磁辐射,加入随机信息等来降低非入侵攻击所能测量到的敏感数据特征;加入开关、电路等对攻击进行检

测，例如，用开关检测电路物理封装是否被打开等。在关键应用如工业控制中还可使用容错硬件设计和可靠性电路设计。

(2) 硬件安全架构层。该方法借鉴了可信平台模块(Trusted Platform Module，TPM)的思路。可采取的措施包括：加入部分硬件处理机制支持加密算法甚至安全协议；使用分离的安全协处理器模块，用来处理所有的敏感信息；使用分离的存储子系统(RAM、ROM、FLASH 等)作为安全存储区域，这种隔离可以限制只有可靠的系统部件才可以对安全存储区域进行存取；如果上述功能不能实现，可以利用存储保护机制，即通过总线监控硬件来区分对安全存储区域的存取是否合法来实现，对经过总线的数据在进入总线前进行加密以防止总线窃听等。实际的例子包括 ARM 公司的 Trustzone 和 Intel 的 LaGrande 等。

(3) 软件安全架构层。该层主要通过增强操作系统或虚拟机(如 Java 虚拟机)的安全性来增强系统安全。例如，微软的 NGSCB(Next-Generation Secure Computing Base，下一代安全计算基础)，通过与相应硬件(如 Intel LaGrande)协同工作提供如下增强机制：进程分离(Process Isolation)，用来隔离应用程序，免受外来攻击；封闭存储(Sealed Storage)，让应用程序安全地存储信息；安全路径(Secure Path)，提供从用户输入到设备输出的安全通道；证书(Attestation)，用来认证软/硬件的可信性。其他方法还有通过加强 Java 虚拟机的安全性，对非可靠的代码使其在受限制和监控的环境中运行(如沙盒 Sand Box)等。另外，该层还对应用层的安全处理提供必要的支持。例如，在操作系统之内或之上充分利用硬件安全架构的硬件处理能力优化和实现加密算法，并向上层提供统一的应用编程接口等。

(4) 安全应用层。通过利用下层提供的安全机制，实现涉及敏感信息的安全应用程序，保障用户数据安全。这种应用程序可以是包含诸如提供 SSL 安全通信协议的复杂应用，也可以是仅仅简单查看敏感信息的小程序，必须符合软件安全架构层的结构和设计要求。

4.1.2　TinyOS 与 TinyECC 简介

上一节从嵌入式系统安全的一般层面来讨论，本节讨论一个传感器结点中常用的操作系统 TinyOS 系统并简要介绍其安全性作为案例分析(Case Study)。其实，由于传感器硬件平台的资源极为有限，典型的嵌入式操作系统如 VxWorks，QNX 等功能过于复杂，可能很难在传感器硬件平台上高效运行。目前已经出现多种适合于无线传感器网络应用的操作系统，如 TinyOS、MantisOS、SOS 等。其中 TinyOS 是目前无线传感器网络研究领域使用最为广泛的操作系统。

TinyOS 是美国加州大学伯克利分校开发的用于传感器结点的开源操作系统。其设计的主要目标是代码量小、耗能少、并发性高、鲁棒性好，可以适应不同的应用。是一种基于组件的(component-based)操作系统。系统由一个调度器和一些组件组成。组件由下到上可分为硬件抽象组件、综合硬件组件和高层软件组件。高层组件向底层组件发出命令，底层组件向高层组件报告事件。调度器具有两层结构，第一层维护着命令和事件，它主要是在硬件中断发生时对组件的状态进行处理；第二层维护着任务(负责各种计算)，只

有当组件状态维护工作完成后，任务才能被调度。TinyOS 的组件层次结构类似于网络协议栈，底层的组件负责接收和发送最原始的数据位，而高层的组件对这些位数据进行编码、解码，更高层的组件则负责数据打包、路由和传输数据。TinyOS 体系结构如图 4.3 所示。

主组件(含调度程序)		
应用(用户组件)		
行动	感知	通信
硬件抽象层		

图 4.3 TinyOS 体系结构

TinyOS 是用 nesC 语言编写的，基于 TinyOS 的应用程序也使用 nesC 语言编写。nesC 是专门为资源极其受限、硬件平台多样化的传感器结点设计的开发语言。TinyOS 的程序核心往往很小(核心代码和数据大概在 400 字节左右)，能够有效缓解传感器存储资源少的问题，使其有效地运行在传感器硬件平台上。它还提供一系列可重用的组件，可以简单方便地编写程序，获取和处理传感器的数据并进行无线传输。这种组件编程的最大优点是可以增加代码的复用性，使得代码更容易移植到不同的结点平台上。而且采用这种基于组件的开发，能够快速实现各种应用。

为了简化设计并降低实现开销，TinyOS 核心使用了事件驱动的单线程任务调度机制，这和传统操作系统的多线程调度机制完全不同。这意味着在任何时刻，处理器只能执行一个任务。因而如果当前正在执行一个任务，处理器必须等这个任务处理完毕，才能开始处理另一个任务。为了保证系统的响应性，一般单个 TinyOS 任务的执行时间不能太长。另外，单个 TinyOS 任务中不能有 IO 等阻塞的调用。

TinyOS 中的通信采用主动消息通信模型，它是一个简单的、可扩展的、面向消息通信的高性能通信模式，一般广泛应用在并行分布式处理系统中。主动消息不但可以让应用程序开发者避免使用忙等待方式等待消息数据的到来，而且可以在通信和计算之间形成重叠，可以加大地提高 CPU 的使用效率，并减少传感器结点的能耗。如果把主动消息通信实现为一个组件，那么就可以屏蔽下层不同的通信硬件，为上层应用提供基本的、一致的通信原语，方便应用层开发人员开发各种不同的应用。

TinyOS 提供了一系列关键服务，包括：

(1) 核心服务，如读取传感器、串口通信、读取程序内存和外存、基本的点对点传输服务等。

(2) 数据收集协议，如 CTP。CTP 集成了链路重传、链路估计等技术，可以将多个结点上的数据通过该多跳路由传到汇聚结点。

(3) 数据分发协议，如 Drip 和 Dip。两者可以通过汇聚结点分发多种系统参数，并在网络内维持一致性。

(4) 时间同步协议，如 FTSP。FTSP 协议通过在网络内交换同步消息，达到全网同步。

(5) 网络重编程协议，如 Deluge。Deluge 协议可以通过汇聚结点分发程序代码，并通过结点自编程，达到应用程序更新以及程序重编程的目的。

在 TinyOS 上可以运行 TinyECC，TinyECC 是北卡州立大学 Peng Ning 教授的团队开发的一款基于 ECC 的软件包，提供了签名方案(ECDSA)，密钥交换协议(ECDH)，以及加密方案(ECIES)。该软件使用了一些优化选项，可以根据开发的实际需要使用这些

选项。

TinyECC 2.0 是用于 TinyOS 2.x 平台的一个版本。使用 nesC 实现，并可以根据特定的传感器平台进行优化，已经在传感器硬件 MICA2/MICAz，TelosB/Tmote Sky，BSNV3，以及 Imote2 上测试过，并且支持 SECG 推荐的 128 位、160 位、192 位椭圆曲线域参数。

4.2　智能手机系统安全

据 Gartner 预计，到 2013 年，全球 PC 保有量将达到 16.2 亿台，而智能手机和具备浏览器的传统手机的保有量将达到 16.9 亿部。智能手机将超越 PC 而成为人们的主要上网工具。因此，移动互联网尤其是智能终端安全将是一个重要的安全课题。

智能手机系统安全主要涉及手机操作系统安全及手机病毒的防治。目前智能手机采用的操作系统主要有：Google Android 平台、苹果的 iOS 系统、微软的 Windows Mobile 操作系统（Window Phone 7，Windows 8）、以 Nokia 为主要发起厂商的 Symbian 操作系统以及 Palm 操作系统、Linux 等。其中应该关注中国移动的 OMS（Open Mobile System）系统，因为国产手机操作系统及其硬件平台的研发具有重要战略意义，为此，工信部于 2012 年开展了核高基重大专项“移动智能终端操作系统研发”。

随着终端操作系统的多样化，手机病毒将呈现多样性的趋势。随着基于 Android 操作系统的智能手机快速发展，基于此种操作系统的手机也日渐成为黑客攻击的目标。因此，在一般性介绍智能手机病毒后，分别介绍 Android 系统和 OMS 系统。

4.2.1　智能手机病毒简介

手机病毒会利用手机操作系统的漏洞进行传播。手机病毒是以手机为感染对象，以通信网络（如移动通信网络、蓝牙、红外线）为传播媒介，通过发送短信、彩信、电子邮件、聊天工具、浏览网站、下载铃声等方式进行传播。手机病毒的主要危害可以分为以下 4 种情况。

（1）导致用户手机里的个人隐私外泄。

（2）控制手机进行强行消费，拨打付费电话，订购高额短信服务，导致通信费用剧增。

（3）通过手机短信的方式传播非法信息，如发送垃圾邮件、垃圾短信等。

（4）破坏手机软件或者硬件系统，如 SIM 卡损毁，造成手机通信瘫痪，如手机死机等。

同计算机病毒类似，手机病毒具有病毒的一般特性。

（1）传播性：手机病毒具有把自身复制到其他设备或者程序的能力，手机病毒可以自我传播，也可将感染的文件作为传染源，并借助该文件的交换、复制再传播，感染更多手机。

（2）隐蔽性：手机病毒隐藏在正常程序中，当用户使用该程序时，病毒乘机窃取系统的控制权，然后执行病毒程序，而这些动作是在用户没有察觉的情况下完成的。

(3) 潜伏性：病毒感染系统后不立即发作，可能在满足触发条件时才开始发作。

(4) 破坏性：无论何种手机病毒，一旦侵入手机都会对手机软/硬件造成不同程度的影响，轻则降低系统性能、破坏丢失数据和文件导致系统崩溃，重则可能损坏硬件。

手机病毒的分类依据包括：工作原理、传播方式、危害对象和软件漏洞出现的位置。

(1) 根据手机病毒的工作原理划分，手机病毒可以分为以下 5 类。

① 引导型病毒：智能手机具有操作系统，引导型病毒是一种在系统开机自检（BIOS）完成后，进入操作系统引导时开始工作的病毒。引导型病毒先于操作系统执行。病毒先获得控制权，将真正的引导区内容转移或替换，待病毒程序执行后，再将控制权交给真正的引导区内容，带病毒的系统看似正常运转，其实病毒已隐藏在系统中。

② 宏病毒：宏病毒是一种寄存在文档或模板（如 Word、PowerPoint 文件等）宏中的病毒，宏是一种可以自动执行的代码，一旦打开这样的文档，其中的宏病毒会执行。宏病毒主要是使用某个应用程序自带的宏编程语言（如 VB Script）编写。智能手机（如 Windows Phone 7）可以安装阅读 Word 和 PowerPoint 文档的应用软件，可能遭到这种病毒的攻击。

③ 文件型病毒：文件型病毒是主要感染可执行文件（如 apk 文件）的病毒，它通常隐藏在宿主（Host）程序中，执行宿主程序时，先执行病毒程序再执行宿主程序。它的安装必须借助病毒的装载程序，已感染病毒的文件执行速度会减慢。

④ 蠕虫（Worm）病毒：蠕虫的特征是在手机和手机之间自动地自我复制，它接管了手机中传输文件或信息的功能。一旦手机感染蠕虫病毒，蠕虫即可独自大量复制和传播。

⑤ 木马（Trojan Horse）病毒：这类病毒是在正常程序中植入恶意代码，当用户启动程序时，该恶意代码也同时运行，并做一些破坏性动作。

(2) 根据手机病毒传播方式划分，手机病毒可以分为 4 类：通过手机外部通信接口进行传播，如蓝牙、红外、Wi-Fi 和 USB 等；通过互联网接入进行传播，如网站浏览、电子邮件、网络游戏、下载程序、聊天工具等；通过电信增值服务（业务）传播，如 SMS、MMS 等；通过手机自带应用程序进行传播，如 Word 文档、电子书等。

(3) 根据手机病毒的危害对象划分：危害手机终端的病毒、危害移动通信核心网络的病毒。

(4) 根据软件漏洞出现的位置划分：手机操作系统漏洞病毒、手机应用软件漏洞病毒、交换机漏洞病毒、服务器漏洞病毒。

安全的手机操作系统通常具有如下 5 种特征。

(1) 身份验证：确保所有访问手机的用户身份真实可信。可以采用的身份认证方式有口令认证、智能卡认证、生物特征识别（如指纹识别）及实体认证机制等方式。

(2) 最小特权：每个用户在通过身份验证后，只拥有恰好能完成其工作的权限，即将其拥有的权限最小化。

(3) 安全审计：对指定操作的错误尝试次数及相关安全事件进行记录、分析。

(4) 安全域隔离：安全域隔离分为物理隔离和逻辑隔离。物理隔离是指对移动终端中的物理存储空间进行划分，不同的存储空间用于存储不同的数据或代码，而逻辑隔离主要包括进程隔离、数据的分类存储。

（5）可信连接：对于无线连接（蓝牙、红外、WLAN 等），默认属性应设为“隐藏”或者“关闭”以防非法连接；在实际连接时，需要对所有请求连接进行身份认证。

4.2.2　Android 系统简介

最后简要介绍一下 Android 手机操作系统。Android 是 Google 与 OHA（Open Handset Alliance，开放手机联盟）合作开发的基于 Linux 2.6 平台的开源智能手机操作系统平台。其系统架构如图 4.4 所示，包括 4 层结构。

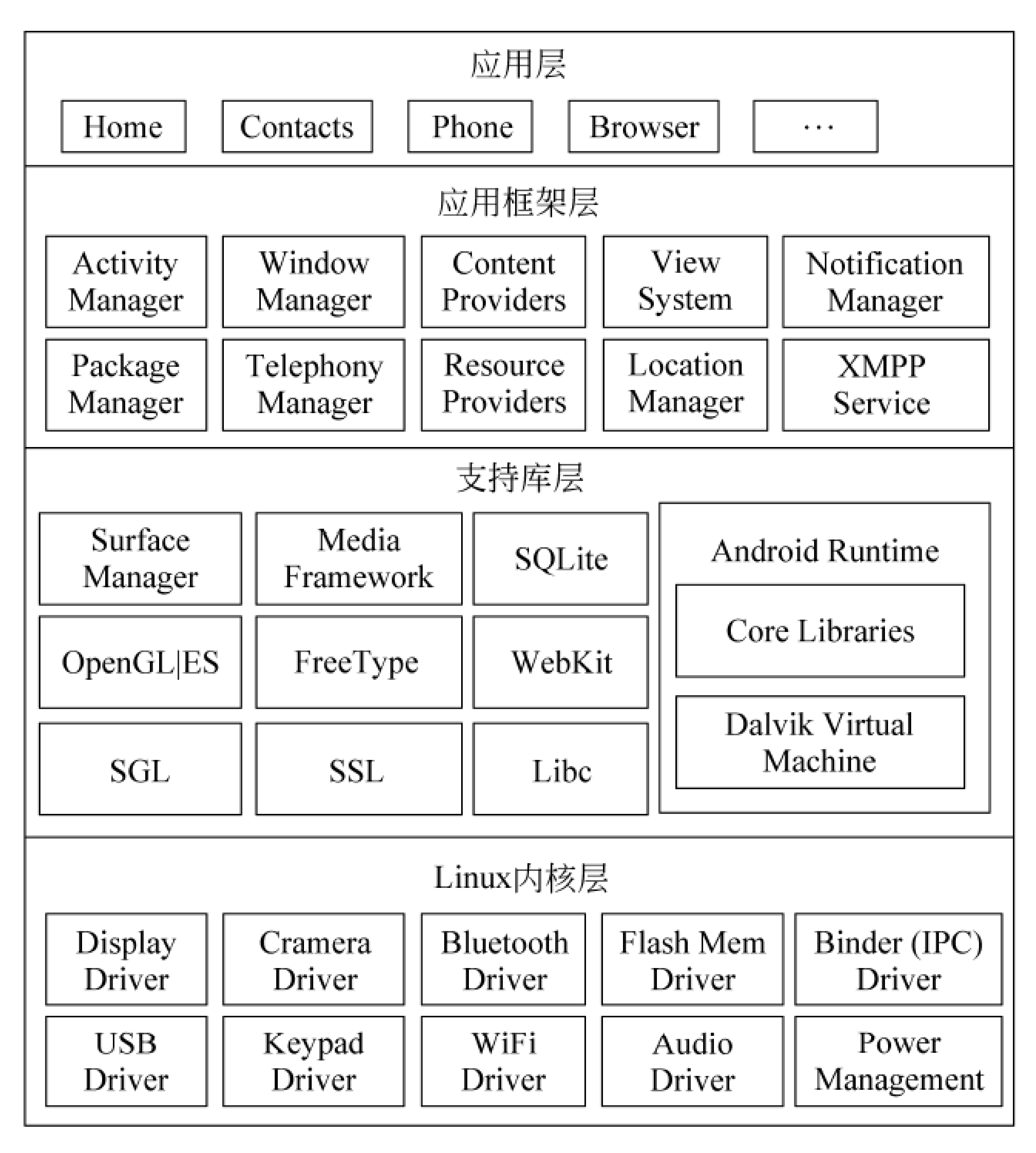


图 4.4　Android 系统架构图

（1）应用层（Applications）。Android 操作系统的用户应用层，直接面向用户，完成显示以及与用户交互的功能，包括一系列主要的应用程序包，如 E-mail 客户端、SMS 短信程序、浏览器等。

（2）应用框架层（Application Framework）。该层专门为应用程序的开发而设计，提供允许开发人员访问核心应用程序所使用的 API 框架。它由一系列的服务和系统构成，提供功能管理和组件重用机制，包含电源管理、窗体管理、资源管理等。

（3）支持库层（Libraries），包含虚拟机（Runtime）。这一层主要与进程运行相关，Dalvik 虚拟机 DVM 是类似 JVM 的虚拟机，提供 Java 语言的运行环境，每一个 Android 程序都有独立的 Dalvik 虚拟机为它提供运行环境。核心库（Core Libraries）提供了 Java 编程语言核心库的大多数功能。库中的代码主要基于 C/C++，为上层的应用程序框架提供访问硬件的方式，可用于较底层的应用程序，其中比较重要的是对 SQL Lite 的支持，

2D/3D 图像技术的支持，以及多媒体解码等。

(4) Linux 内核层(Linux Kernel)。Android 的内核为 Linux 2.6 内核，Linux 内核为 Android 手机提供了一系列硬件驱动，它主要用于保障安全性、内存管理、进程管理、网络协议栈等。

*4.2.3 OMS 平台简介

OMS(Open Mobile System)是中国移动通讯集团基于 Google Android 1.5 平台设计的一种更适合中国国情和中国人习惯的智能手机操作系统，OMS 系统与 Android 系统同样采用 Linux 内核，通过 TD 通信模块以 Modem AP 的方式桥接使 Android 平台兼容中国移动 TD-SCDMA 网络，第一款搭载 OMS 系统的手机是联想 O1。

OMS 是一个开放的移动互联终端软件平台，包括一个 Linux 操作系统，一个 dalvik 虚拟机，一个 Web 浏览器，中间件和一些关键应用。OMS 来源于 Android 平台，除了包括了 Android 的组成部分外，OMS 集成了很多中间件，以及中国移动的增值服务。

在移动业务层面，OMS 内置了包括飞信、快讯、无线音乐随身听、139 邮箱、移动梦网、号簿管家、百宝箱等业务。在手机基本功能层面，OMS 更符合中国人的操作习惯，如对话模式和文件夹模式可以随意选择的短信息用户界面，还有其可以随意定制的主屏幕，奇妙的解锁方式等都体现了 OMS 系统对手机基本功能的重视。在用户体验层面，OMS 系统参考了所有的智能平台的诸多优势，接合中国人的习惯设计出与 Android 系统完全不相符的界面 UI，不但切换主屏是特效炫目，而且实用性上更上一层楼。

OMS 来源于 Android 平台，因此 OMS 系统与 Android 系统在程序兼容性上完全不存在问题。Android Market 提供了丰富的扩展应用，大部分的程序几乎不需任何改动即可直接应用在 OMS 平台上。OMS 终端总体架构如图 4.5 所示。

OMS SDK 是支持两类应用程序的开发：OMS 应用程序和 Widget 应用程序。

OMS 应用程序是基于 Java 的应用程序，类似于 Android 应用程序。但是与 Android 相比，OMS 提供许多特有的 OMS API。可以使用 OMS API 和 Android API 开发 OMS 应用程序。任何使用 Android API 创建的应用程序可以正常运行在 OMS 手机上，然而使用了 OMS 扩展 API 的应用程序不能在 Android 手机上运行，这些 API 需要 OMS 平台的高级特征。可以在 Eclipse IDE 里使用 Android 开发工具(ADT)创建 OMS 应用程序。ADT 插件包括多种强大的扩展，使得创建、编辑、运行和调试 OMS 应用程序更快更方便。

OMS 支持的第二类应用程序是 Widget 应用程序(如 XHTML、CSS、JavaScript 等)。Widget 应用程序是 OMS 的亮点，因为 Android 还不支持 Widget 应用程序开发。在 OMS 中，Widget 应用程序遵循 JIL(Joint Innovation Lab)Widget 规范。若使用 Eclipse IDE 开发 Widget 应用程序，可安装一个定制插件 WDT(Widget Development Tools)，该插件集成了对 Widget 项目的支持。WDT 插件包多个功能强大的扩展，使得创建、编辑、构建、运行和调试 Widget 应用程序更加快捷方便。

OMS 作为国内企业主导开放的移动终端操作系统，虽然在推广上面临诸多的挑战，但从系统安全和国家安全的角度而言，这样做是有必要的。目前关于 OMS 平台的安全

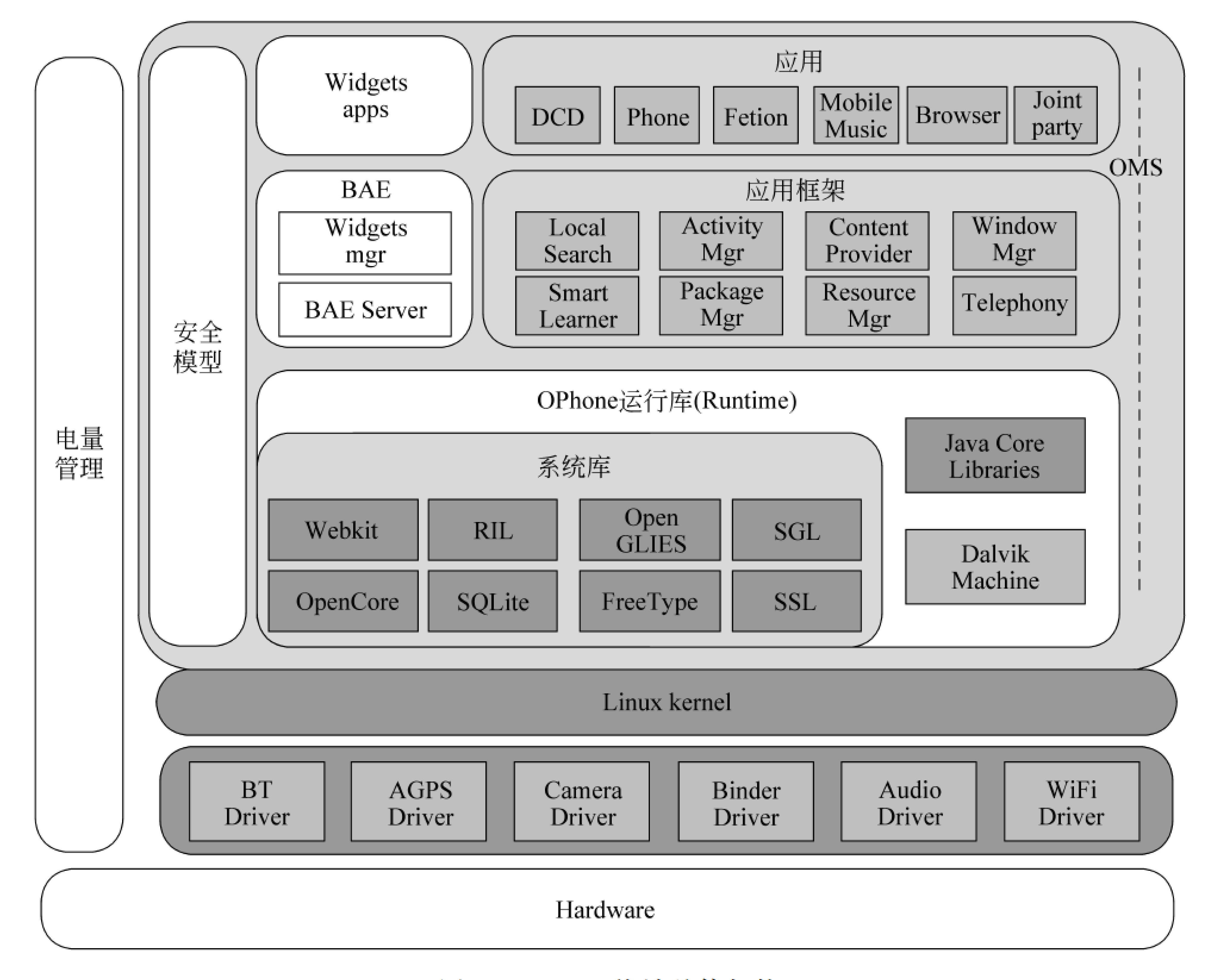


图 4.5　OMS 终端总体架构

性分析方面的文献，仍然十分少见。据报道，OMS 在多个层面引入安全策略以保证移动终端和用户数据的安全性，同时，还具备系统备份还原机制防止用户数据丢失。

随着移动互联网大潮的来临，以及智能手机以及平板电脑的普及，智能手机应用以及平板电脑应用将逐步增多。智能手机的平台安全和应用安全均将会越来越受到重视。

研究与思考

[1]　分析苹果 iOS 系统以及 Android 系统的安全性。比较两者的安全性差异。

[2]　分析 OMS 平台的安全性。

[3]　在 OMS 平台实现一个手机流量监控的安全软件。

[4]　研发一种工具可以进行智能手机病毒的分析与防范。

[5]　研发一种工具用于智能手机（如 iOS/Android 平台）的计算机犯罪取证。

进一步阅读建议

近年来对国际计算机安全界对 Android 系统安全的关注逐步增多。

[1] A. Shabtai, Y. Fledel, U. Kanonov, Y. Elovici, S. Dolev, C. Glezer, Google Android: A Comprehensive Security Assessment [J], IEEE Security & Privacy, 8(2):35-55, 2010.
[2] A. Shabtai, Y. Fledel, Y. Elovici, Securing Android-Powered Mobile Devices Using SELinux [J], IEEE Security & Privacy, 8(3):36-44, 2010.
[3] D. Barrera, P. Van Oorschot, Secure Software Installation on Smartphones [J], IEEE Security & Privacy, 9(3):42-48, 2011.
[4] C. Miller, Mobile Attacks and Defense [J], IEEE Security & Privacy, 9(4):68-70, 2011.
[5] D. Barrera, H. Kayacik, P. C. Oorschot, A. Somayaji, A Methodology For Empirical Analysis Of Permission-Based Security Models And Its Application to Android[C], In Proc. of the 17th ACM conference on Computer and Communications Security (CCS'10) , October 2010, 73-84.

本章参考文献

[1] 苹果公司已申请 RFID 专利用于下一代 iPhone: http://www.eeworld.com.cn/xfdz/2011/0421/article_6205.html.
[2] 郭春霞,裘雪红. 嵌入式系统安全的研究与设计[J]. 电子科技,2008(5).
[3] 刘云浩. 物联网导论[M]. 北京:科学出版社,2011.
[4] TinyECC: A Configurable Library for Elliptic Curve Cryptography in Wireless Sensor Networks (Version 2.0), http://discovery.csc.ncsu.edu/software/TinyECC/.
[5] 王明剑等. 嵌入式信息系统安全体系构建与应用[J]. 信息安全与通信保密,2005(7).
[6] 落红卫. 手机病毒及应对技术探究[J]. 信息网络安全,2009(9).
[7] 肖云鹏,刘宴兵,杨莎莎,徐光侠,基于云计算的 OMS 即时通讯系统设计与实现[J]. 重庆邮电大学学报(自然科学版),22(4),2010 年 8 月,468-472.
[8] Wei Ren, Yuliang Liu, Junge Zhao, Provably Secure Information Hiding via Short Text in Social Networks, Tsinghua Science and Technology, Vol. 17, No. 3, 2012.
[9] 任伟,软件安全[M]. 北京:国防工业出版社,2010.
[10] Wei Ren, Yuliang Liu, Junge Zhao, Provably Secure Information Hiding via short Text in social Networking Tools, Tsinghua science & Technology, Vol. 17, No. 3. 2012.

第 2 部分　物联网网络层安全

第 5 章　近距离无线接入安全——无线局域网安全

物联网中的感知层终端系统，如 RFID 读写器、无线传感器网络的网关结点、以及智能手机都可以通过无线局域网连接到 Internet，因此，需要考虑无线局域网的安全。

早期版本的 IEEE 802.11 无线局域网（Wireless Local Area Network，WLAN）有一个特定的安全架构，称为 WEP（Wired Equivalent Privacy，有线等效保密）。其含义是 WLAN 至少要和 LAN（即有线局域网）的安全性相当（等价）。例如，一个攻击者希望连接一个 LAN，需要物理上接入集线器，然而集线器通常锁在房间或柜子里，所以很难办到。但是对 WLAN 而言攻击者就很容易，因为此时接入网络不需要从物理上接入任何设备，设计 WEP 的目的之一便是设法阻止这种非授权的接入。总的来说，WEP 要使得攻击 WLAN 的难度与攻击 LAN 的难度相当，除阻止非授权的接入外，还包括阻止对通信消息的窃听与破坏。但实际上 WEP 没有达到这一目的。为了改进 WEP 的安全性，IEEE 后来提出了 WLAN 的一种新的安全架构，称为 IEEE 802.11i。同时，我国提出了针对 WLAN 安全的国际标准 WAPI。本章将分别加以介绍。

5.1　无线局域网的安全威胁

5.1.1　无线局域网的网络结构

WLAN 的工作模式可分为基础结构（Infrastructure）模式和自组织（Ad Hoc）模式，基础结构的拓扑结构是扩展服务集（Extended Service Set，ESS），而自组织网络的拓扑结构是独立基本服务集（Independent Basic Service Set，IBSS）。在基础结构网络下，无线终端（Station，STA）通过访问结点（Access Point，AP）相互通信，而且可以访问有线网络，是最常用的网络拓扑结构；自组织网络是无线终端 STA 之间相互连接通信形成的一种工作方式。

1. 基础结构无线局域网

基础结构模式的无线局域网中，所有 STA 与 AP 通信，AP 往往还充当网桥的作用，将数据转发到相应的有线或无线网络中，即 STA 通过 AP 实现与 STA 间的通信，或 STA 通过 AP 实现与有线网络的通信。

一个 AP 及若干 STA 组成的通信区域称为一个基本服务集（BSS）。在使用 AP 的 WLAN 中，一个 AP 覆盖的无线连接网络区域为一个 BSS，通过 AP 可以使无线网络连接到有线网络中，扩展连接不同的 BSS 使它们能够相互通信的架构称为分布系统（Distribution System，DS），通过 DS（通常是有线网络）实现 BSS 之间互相通信，这种扩展

的 BSS 称为扩展的基本服务集(EBSS)。上述这种无线网络部署称为基础结构无线局域网,如图 5.1 所示。

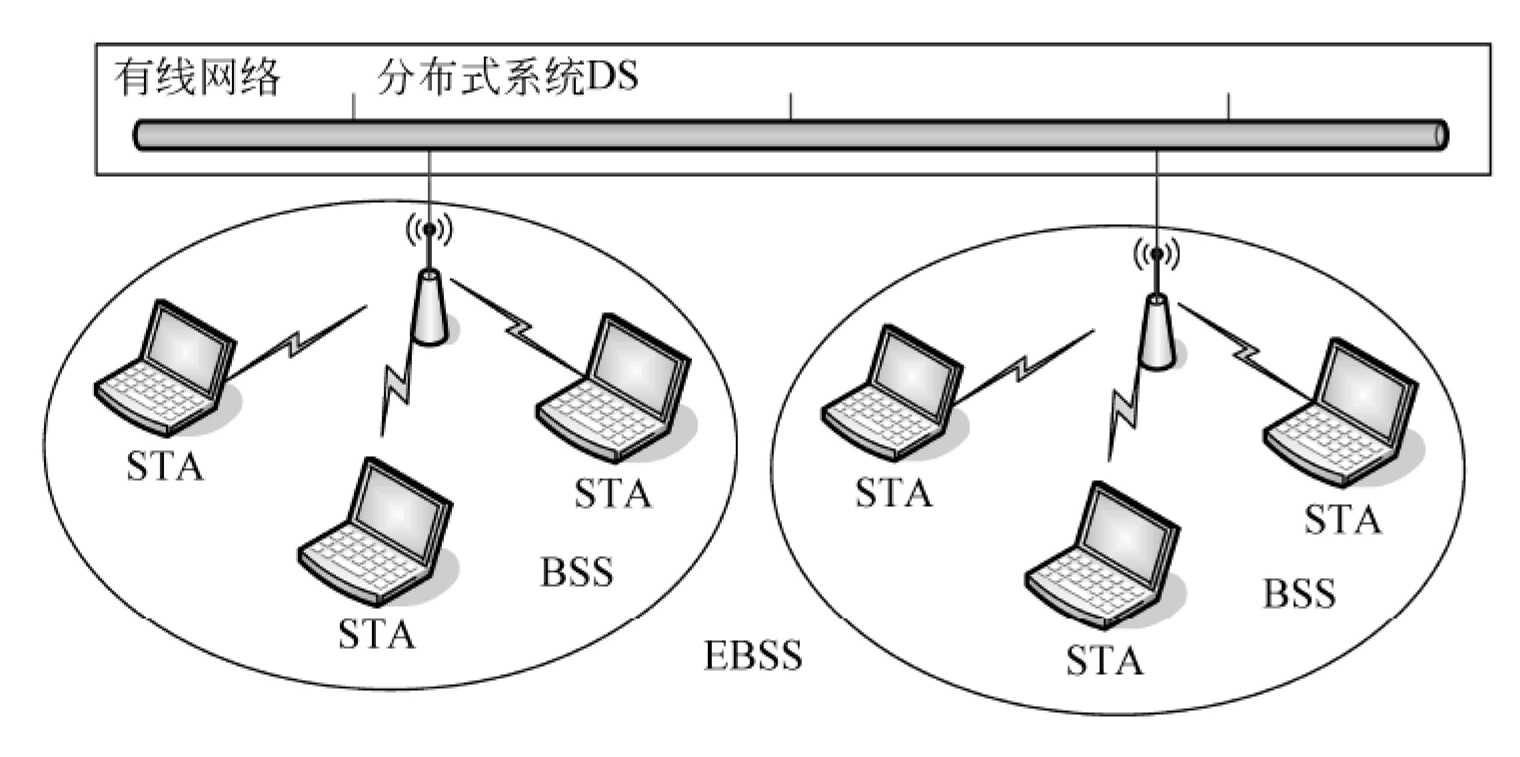


图 5.1　基础结构无线局域网

2. 自组织无线局域网

自组织无线局域网络,也称 Ad Hoc 无线局域网,其拓扑结构为 IBSS,如图 5.2 所示,适用于不存在有线网络的少量主机组建临时性网络。在这种架构中,主机彼此之间直接通信,实际的应用包括野外作业时的资源共享、举办临时流动会议、抢险救灾临时网络等。

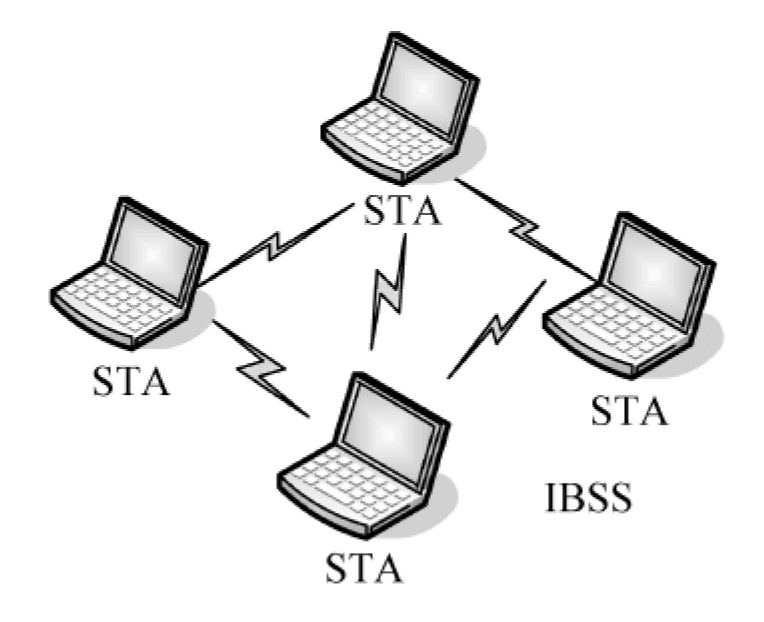


图 5.2　自组织无线局域网络

5.1.2　无线局域网的安全威胁

由于无线局域网通过无线电波传递信息,所以在数据发射机覆盖区域内的几乎任何一个 WLAN 用户都能接触到这些数据。WLAN 所面临的基本安全威胁主要有信息泄漏、完整性破坏、拒绝服务和非法使用。主要的威胁包括非授权访问、窃听、伪装、篡改信息、否认、重放、重路由、错误路由、删除消息、网络泛洪等,可见,均为常见的无线网络威胁。

(1) 非授权访问:入侵者访问未授权的资源或使用未授权的服务。入侵者可查看、删除或修改未授权访问的机密信息,造成信息泄漏、完整性破坏,以及非法访问和使用资源。

(2) 窃听:入侵者能够通过通信信道来获取信息。AP 的无线电波难以精确地控制在某个范围之内,所以在 AP 覆盖区域内的几乎任何一个 STA 都能够窃听这些数据。

(3) 伪装:入侵者能够伪装成其他 STA 或授权用户,对机密信息进行访问;或者伪装成 AP,接收合法用户的信息。

(4) 篡改信息:当非授权用户访问系统资源时,会篡改信息,从而破坏信息的完整性。

(5) 否认:接受信息或服务的一方事后否认曾经发送过请求或接收过该信息或服务。这种安全威胁通常来自系统内的合法用户,而不是来自未知的攻击者。

(6) 重放、重路由、错误路由、删除消息：重放攻击是攻击者复制有效的消息事后重新发送或重用这些消息以访问某种资源；重路由攻击(主要是在 Ad Hoc 模式中)是指攻击者改变消息路由以便捕获有关信息；错误路由攻击能够将消息路由到错误的目的地；而删除消息是攻击者在消息到达目的地前将消息删除掉，使得接收者无法收到消息。

(7) 网络泛洪：入侵者发送大量伪造的或无关消息从而使得 AP(或者 STA)忙于处理这些消息而耗尽信道资源和系统资源，进而无法对合法用户提供服务。

5.2　无线局域网的安全机制

5.2.1　WEP 加密和认证机制

1. WEP 加密

在 IEEE 802.11 1999 年版本的协议中，规定了安全机制 WEP[1]，WEP 提供 3 个方面的安全保护：数据机密性、数据完整性，以及认证机制(本节介绍前两者，后者下一节介绍)。其中使用了 RC4 序列密码算法[2,3]，用密钥作为种子通过 RC4 算法产生伪随机密钥序列(PRKS)，然后和明文数据异或后得到密文序列。WEP 协议加密流程如图 5.3 所示。

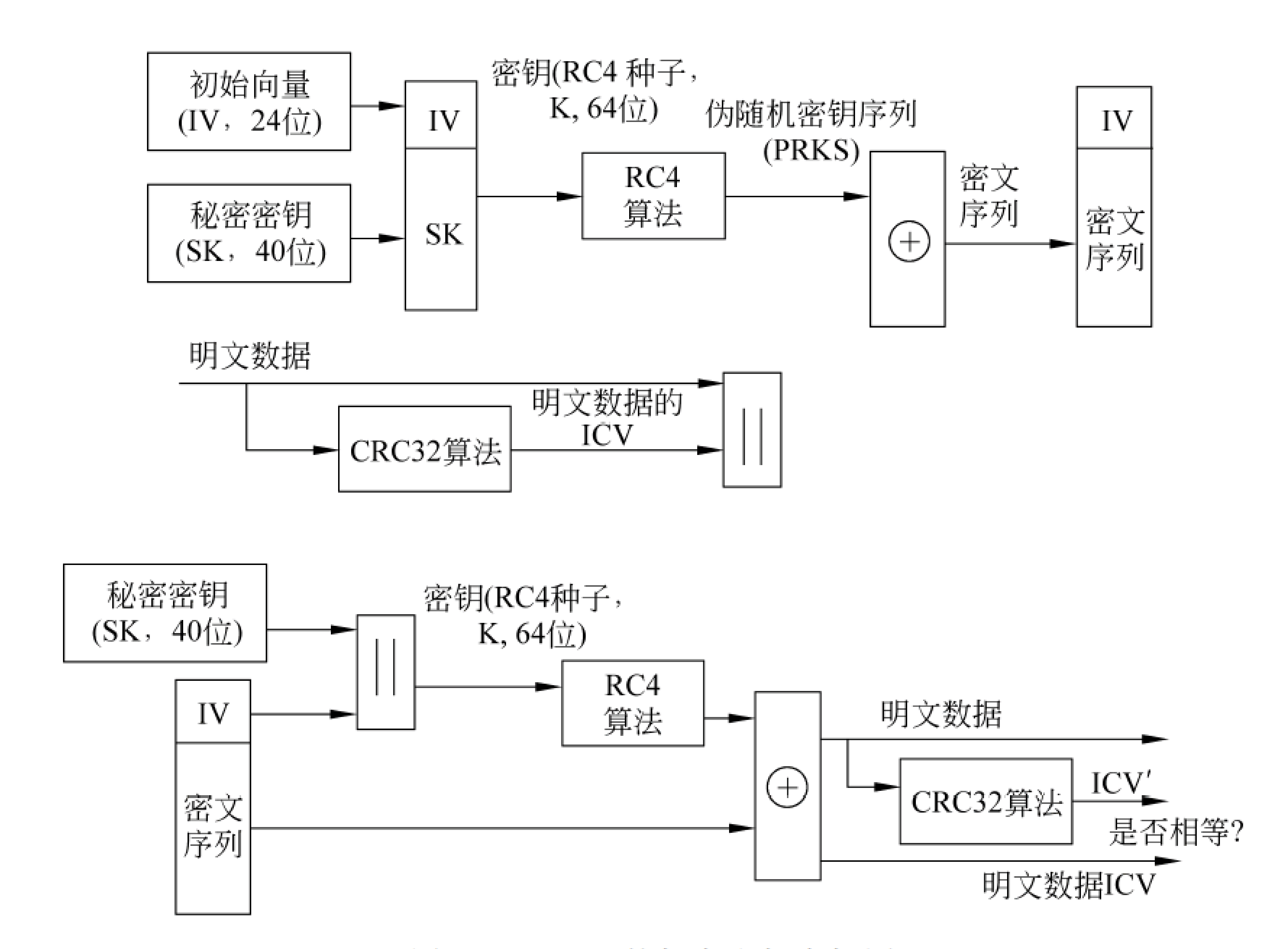


图 5.3　WEP 的加密和解密框图

由于在序列密码算法中，同一伪随机密钥序列不能使用两次，因此 WEP 中将 RC4 的输入密钥 K 分为两部分：24 位的初始向量(Initialization Vector，IV)和 40 位的秘密密钥(SK)，IV ‖ SK＝K，每加密一次 IV 需改变一次，IV 以明文的形式随着密文数据帧一起发往接收方。SK 为 BSS 中各 STA 所共享的秘密信息，通常由管理员手工配置和分发。另

外，为了保证数据的完整性，WEP 中采用 CRC32 算法作为消息鉴别算法，并将数据的消息鉴别码 ICV 连同明文数据一起加密得到密文序列，然后发送给接收端。接收端解密密文序列，重新计算消息认证码 ICV′，并和收到的 ICV 比较，若不相同则抛弃接收的密文序列。

WEP 协议希望能提供给用户与有线网络相等价的安全性。然而研究分析[4]表明，WEP 机制存在较大安全漏洞。

(1) WEP 加密是 AP 的可选功能，在大多数的实际产品中(如无线路由器)默认为关闭，因此用户数据还是暴露在攻击者面前。

(2) WEP 对 RC4 的使用方式不正确，易受 IV 弱点攻击，从而破解秘密密钥(SK)。RC4 中存在弱密钥[2,3]。弱密钥是指：RC4 输入该密钥产生的输出伪随机性差，容易出现重复。当使用一个弱密钥作为种子输入到 RC4 时，RC4 输出的前面几位可以推断出其他位。因此，建议抛弃 RC4 输出的前 256 位。

(3) 初始向量(IV)空间太小。序列加密算法的一个重要缺陷是加密使用的伪随机密钥序列不能出现重复。就 WEP 加密而言，如果使用相同的 IV ‖ SK 加密两个消息，攻击者有：

$$C1 \oplus C2 = \{P1 \oplus RC4(IV \parallel SK)\} \oplus \{P2 \oplus RC4(IV \parallel SK)\} = P1 \oplus P2$$

如果其中的一个消息明文已知，另一个消息明文就可以立即获得(这个在密码学中是已知明文攻击中的一种情形)。为了防止这种攻击，WEP 采用 IV-SK 作为密钥，其中 SK 不变，IV 每传输一次改变一次，以获得不同的伪随机密钥序列 PRKS，IV 以明文的方式传送。但是 WEP 协议中的 IV 空间只有 24 位，在实际的产品中 IV 一般用计数器实现，24 位的空间使得在繁忙的 WLAN 中每过几个小时 IV 就会循环重复出现一次(IV 只有 24 位，这意味着只有大约 17×10^6 种可能的 IV 值。一个 WLAN 设备每秒可以传送大约 500 个完整帧，因此，仅需几小时整个 IV 空间便被用完，另外同一 WLAN 中多个设备在不同的 IV 下使用相同的密钥，使得 IV 空间消耗得更快)。一旦所有 IV 都被使用，就要开始重复使用，而重复使用 IV 意味着使用同一个伪随机密钥序列进行加密(因为 PRKS ＝RC4(IV ‖ SK)，而 SK 是不变的)，这在序列密码算法中是不允许的。

(4) WEP 中的 CRC32 算法原本用于检查通信中的随机误码，不具有抗恶意攻击所需要的消息鉴别功能。首先，CRC32 算法是一个线性函数，因此攻击者可以修改密文而不被发现。其次，由于 CRC32 算法是一个不需要密钥的函数，任何知道消息的人都可以自己计算 ICV。如果攻击者获得一个传输帧对应的明文，就可以在无线网络中传输任意的数据。方法如下：令加密消息为 $(M \parallel CRC(M)) \oplus PRKS$，其中 M 为消息，PRKS 是一个伪随机密钥序列，由 IV 和密钥通过 RC4 算法计算得到。CRC(·)表示 CRC 函数，‖ 表示连接。CRC 关于 XOR 运算是线性的，即有 $CRC(X \oplus Y) = CRC(X) \oplus CRC(Y)$。于是在不知道 M 的情况下，攻击者可以修改或伪造消息。以消息的变化为 ΔM 的情况为例，攻击者可以从窃听的 $(M \parallel CRC(M)) \oplus PRKS$，在不知道 PRKS 的情况下，得到消息修改后的密文 $((M \oplus \Delta M) \parallel CRC(M \oplus \Delta M)) \oplus PRKS$。具体步骤是：首先计算 CRC(ΔM)，然后将 $\Delta M \parallel CRC(\Delta M)$ 与窃听的消息 $((M \parallel CRC(M)) \oplus PRKS)$ 异或。因为 CRC(ΔM) 的计算无需密钥，且 CRC 是线性的，因而尽管有加密和 ICV 机制，攻击者也

可以成功地修改消息,如下:

$$((M \| CRC(M)) \oplus PRKS) \oplus (\Delta M \| CRC(\Delta M))$$
$$= ((M \oplus \Delta M) \| (CRC(M) \oplus CRC(\Delta M))) \oplus PRKS$$
$$= ((M \oplus \Delta M) \| CRC(M \oplus \Delta M)) \oplus PRKS$$

2. WEP 认证机制

WEP(即 IEEE 802.11 中的安全机制)认证技术可用于独立基本服务集中的 STA 之间的认证,也可用于基本服务集中的 STA 和 AP 之间的认证。WEP 有两种认证方式:开放系统认证和共享密钥认证。开放系统认证方式实际上没有认证,仅验证标识,即只要 STA 和 AP 的 SSID 是一致的即可,是一种最简单的情况,也是默认方式。

共享密钥认证方式基于密码学安全协议中的“挑战-应答(Challenge-Response)”模式,其基础是基于对称密码学的两方单向认证协议,假定 AP 和 STA 间通过一个独立于 802.11 的安全通道具有一个共享秘密(即共享密钥)。协议过程如图 5.4 所示。

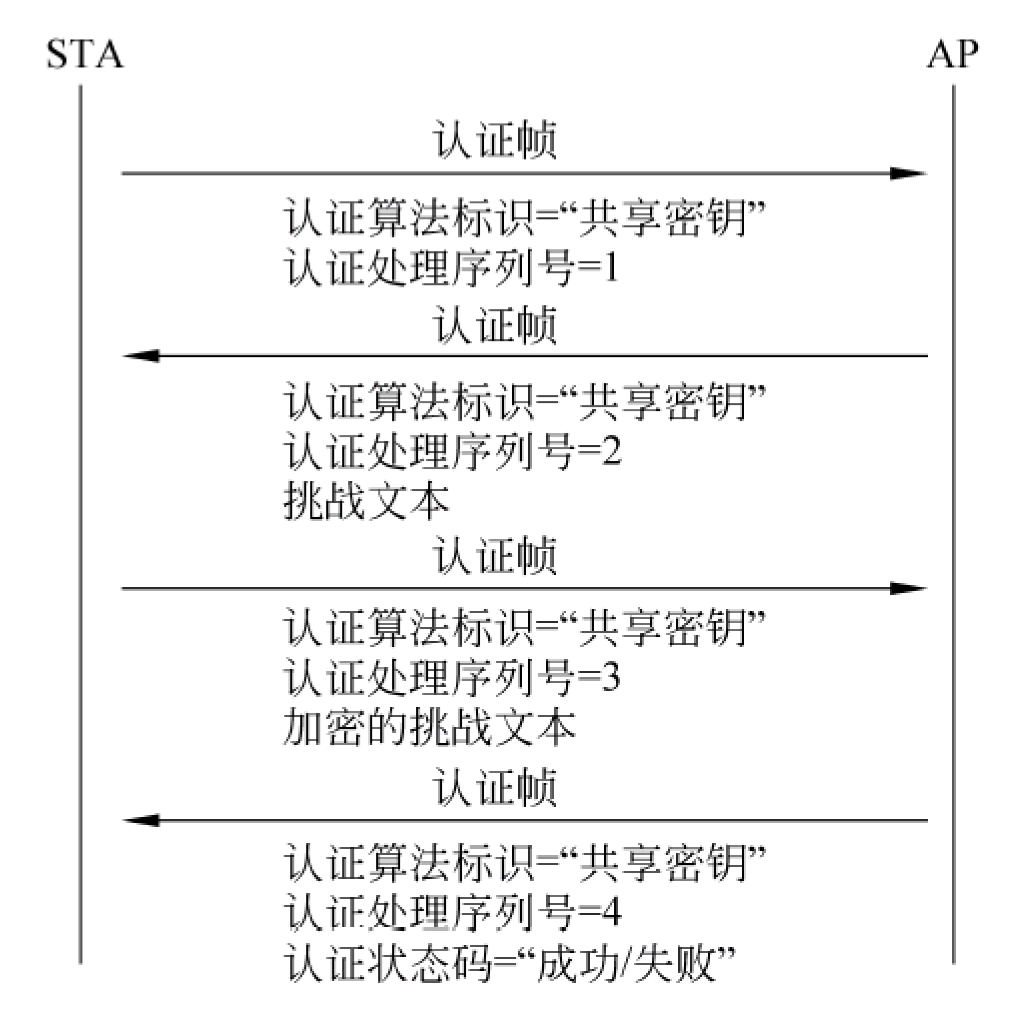


图 5.4 802.11 WEP 共享密钥认证

(1) STA 发送认证帧。

(2) AP 收到后,返回一个认证帧,其帧体包括:认证算法标识=“共享密钥”、认证处理序列号=2、认证状态码=“成功”、认证算法依赖信息=“挑战文本”,如果认证状态码是其他状态,则表明认证失败(例如根据 MAC 地址访问控制列表认为 STA 的 MAC 地址非法),而挑战文本也将不会发送,整个认证过程就此结束。

(3) 如果第(2)步中的状态码=“成功”,则 STA 将从该帧中获得挑战文本并用共享密钥将其加密,然后发送一个认证帧。其帧体包括:认证算法标识=“共享密钥”、认证处理序列号=3、认证算法依赖信息=“加密的挑战文本”。

(4) AP 在接收到第三个帧后,使用共享密钥对加密的挑战文本解密,若和自己发送的相同,则对 STA 的认证成功,否则认证失败。同时 AP 发送一个认证帧,其帧体包括:认证算法标识=“共享密钥”、认证处理序列号=4、认证状态码=“成功/失败”。

WEP 认证机制存在如下问题。

(1) 身份认证是单向的,即 AP 对申请接入的 STA 进行身份认证,而 STA 不能对 AP 的身份进行认证。因此,这种单向认证方式导致有可能存在假冒的 AP。

(2) 从 WEP 协议身份认证过程可以发现,由于 AP 会以明文的形式把挑战文本发给 STA,所以如果能够监听(如利用 Airsnort/BSDAirtools 等工具)一个成功的 STA 与 AP 之间身份验证的全过程,截获它们之间双方互相发送的数据包(挑战文本与加密的挑战文本),就可以计算出用于加密挑战文本的密钥序列。拥有了该密钥序列,攻击者可以向 AP 提出访问请求,并利用该密钥序列加密挑战文本通过认证。

5.2.2 IEEE 802.1X 认证机制

从上节可知 IEEE 802.11 WEP 协议的认证机制存在安全隐患，为了解决无线局域网用户的接入认证问题，IEEE 工作组于 2001 年公布了 802.1X 协议（最新为 2010 版本）。IEEE 802.1X 协议[5]称为基于端口的访问控制协议（Port Based Network Access Control Protocol），它提供访问控制、用户认证，以及计费功能。IEEE 802.1X 本身并不提供实际的认证机制，需要和上层认证协议（EAP）配合来实现用户认证。IEEE 802.1X 在无线网络（WLAN）和有线网络（LAN）中均可应用，其核心是扩展认证协议（Extensible Authentication Protocol，EAP）[6]。图 5.5 给出了 802.1X 协议与上述协议之间的关系。

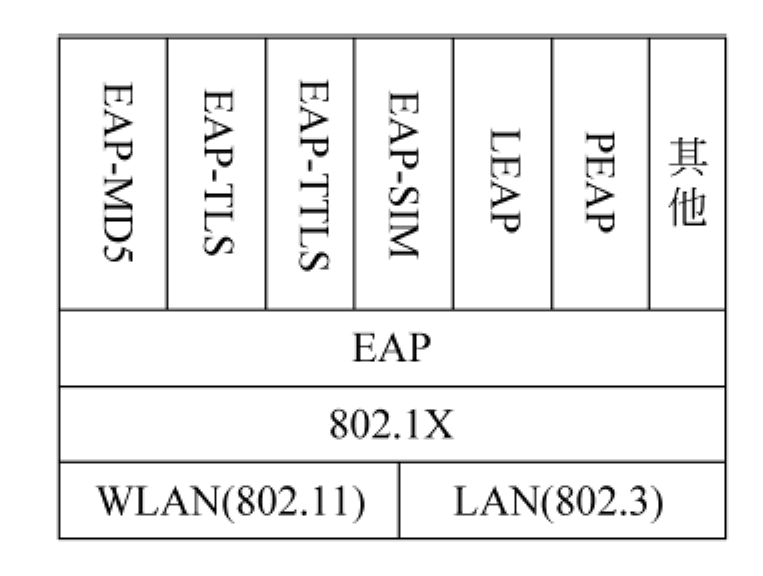


图 5.5　802.1X 协议的位置和 EAP 组成

802.1X 协议是基于 Client/Server 结构的访问控制和认证协议。它可以限制未经授权的用户/设备通过接入端口（Access Port）访问 LAN（或者 AP 访问 WLAN）。在获得交换机或 LAN 提供的各种业务之前，802.1X 对连接到交换机端口上的用户/设备进行认证。在认证通过之前，802.1X 只允许 EAPoL（基于局域网的扩展认证协议，对于 WLAN 情形是 EAPoWLAN）数据通过设备连接的交换机端口；认证通过以后，正常的数据可以顺利地通过以太网端口（或 WLAN 的 AP）。

802.1X 的过程可简单描述为：请求者提供凭证，如用户名/密码、数字证书等，给认证者，认证者将这些凭证转发给认证服务器，认证服务器决定凭证是否有效，并依次决定请求者是否可以访问网络资源，如图 5.6 所示。

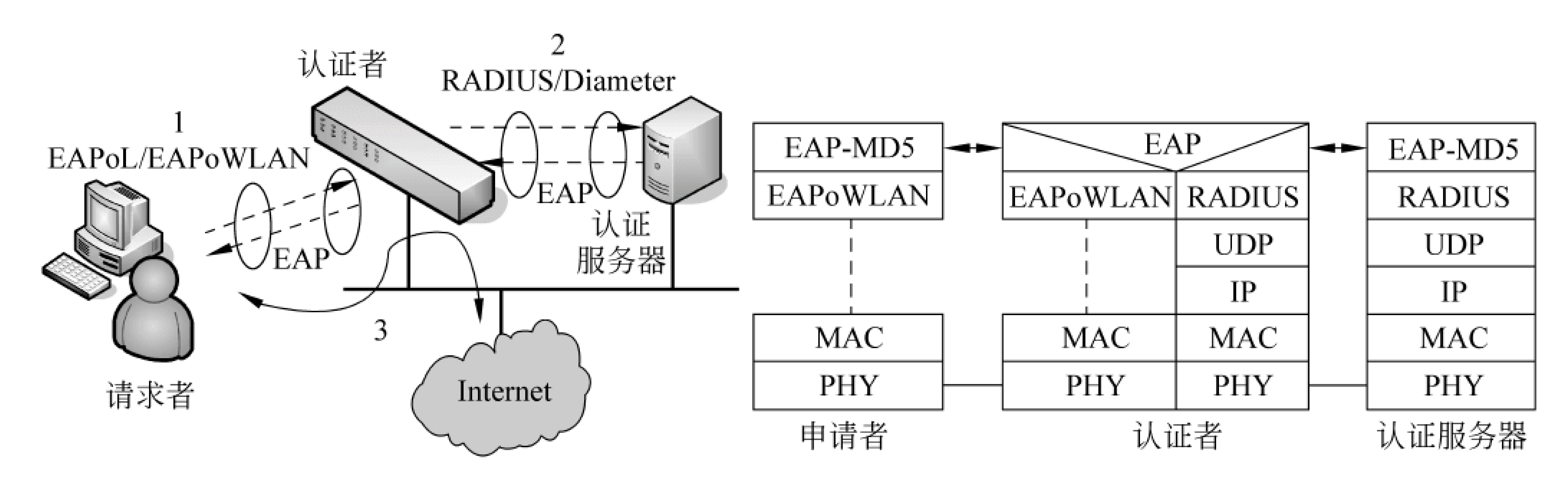


图 5.6　802.1X 的工作过程示意图（以 EAP-MD5 为例）

1. IEEE 802.1X 认证的体系结构

IEEE 802.1X 协议起初是针对以太网提出的基于端口进行网络访问控制的安全标准。基于端口的网络访问控制指的是利用物理层对连接到局域网端口的设备进行身份认证。如果认证成功，则允许该设备访问局域网资源，否则禁止。虽然 802.1X 标准最初是为局域网设计的，后来发现它也适用于符合 802.11 标准的 WLAN，于是被视为是无线局域网增强网络安全的一种解决方案。802.1X 认证的体系结构如图 5.7 所示。

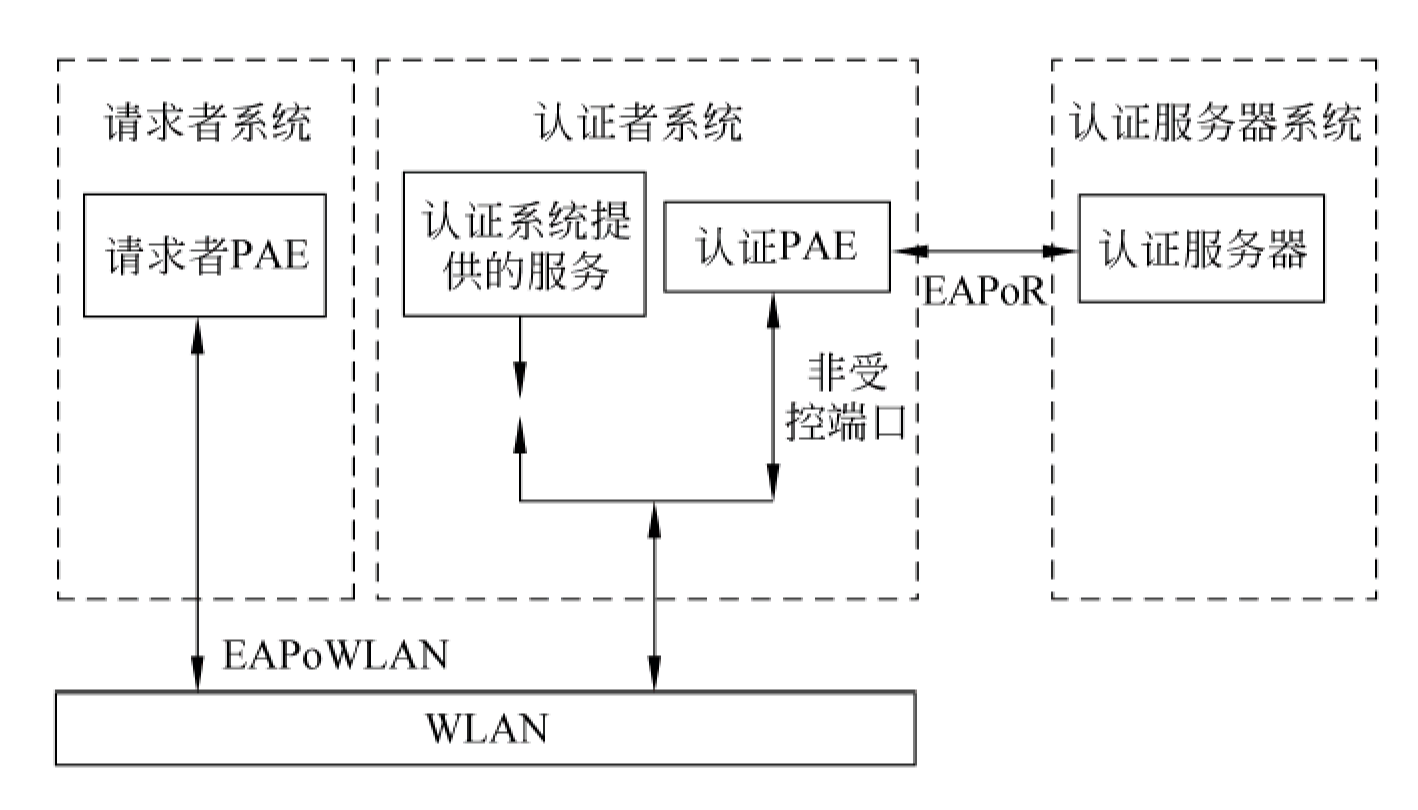


图 5.7　802.1X 认证体系结构

IEEE 802.1X 认证的体系结构从安全协议的角度可视为包括 3 个实体：请求者系统(Supplicant System)、认证者系统(Authenticator System)和认证服务器系统(Authentication Server System)。

从网络的角度则称网络访问的核心部分是 PAE(Port Access Entity，端口访问实体)。在整个认证(访问控制)流程中，端口访问实体包含 3 部分：认证者，即对接入的用户/设备进行认证的端口；请求者，即被认证的用户/设备；认证服务器，即根据认证者的信息，对请求访问网络资源的用户/设备进行实际认证功能的设备。非正式地说，认证者其实是请求者和认证服务器之间的中介。

(1) 请求者系统(也称为客户端系统)：一般为一个用户终端系统(如笔记本电脑)，安装有一个客户端软件，用户通过启动这个客户端软件发起 IEEE 802.1X 协议的认证过程。为了支持基于端口的接入控制，请求者系统必须支持基于局域网(本节指 WLAN)的扩展认证协议(Extensible Authentication Protocol Over LAN，EAPoL)，本节指 EAPoWLAN。

(2) 认证者系统(也称为认证系统)：在无线局域网中就是无线接入点 AP(局域网中是交换机)，是支持 IEEE 802.1X 协议的网络设备。在认证过程中只起到“转发”的功能，所有实质性认证工作在请求者和认证服务器系统上完成。

(3) 认证服务器系统：为认证者提供认证服务的实体，多采用远程身份验证拨入用户服务(Remote Authentication Dial In User Service，RADIUS)[7]。认证服务器对请求方进行认证，然后通知认证者系统这个请求者是否为授权用户。

IEEE 802.1X 认证协议是一种基于端口的对请求者进行认证的方法和策略。通常将物理端口分为两个虚拟端口：非受控端口(Uncontrolled Port)和受控端口(Controlled Port)。非受控端口始终处于双向连通状态(开放状态)，主要用来传递认证信息，如 EAPoWLAN 协议帧(即把 EAP 包封装在 WLAN 上)，可保证随时接收请求者发出的认证请求报文。受控端口受控端口的联通或断开是由该端口的授权状态决定的。认证者的 PAE 根据认证服务器认证过程的结果，控制受控端口的状态：授权(认证、开放)状态或者未授权(未认证、关闭)状态。受控端口平时处于关闭状态，只有在请求者通过认证后才打开，为通过认证验证的用户传递数据和提供服务。如果请求者未通过认证，则受控端口

处于未认证(即关闭)状态,那么请求者无法访问网络服务和资源。通过受控端口与非受控端口的划分,分离了认证数据和业务数据,提高了系统的接入管理和接入服务的工作效率。

在认证时请求者通过非受控端口和 AP(认证者)交互数据,请求者和认证者之间传送 EAPoWLAN 协议帧,认证者和认证服务器同样运行 EAP 协议,认证者将 EAP 封装到其他高层协议中(如 RADIUS)以便 EAP 协议穿越复杂的网络到达认证服务器,称为 EAPoverRADIUS(EAPoR)。若请求者通过了认证,则 AP 为请求者打开一个受控端口,请求者可通过受控端口传输各种类型的数据帧(如 HTTP、POP3)。

2. IEEE 802.1X 协议的认证过程

IEEE 802.1X 协议实际上一个可扩展的认证框架,并没有规定具体的认证协议,具体采用什么认证协议可由用户自行配置,因此具有较好的灵活性。IEEE 802.1X 认证过程如图 5.8 所示。

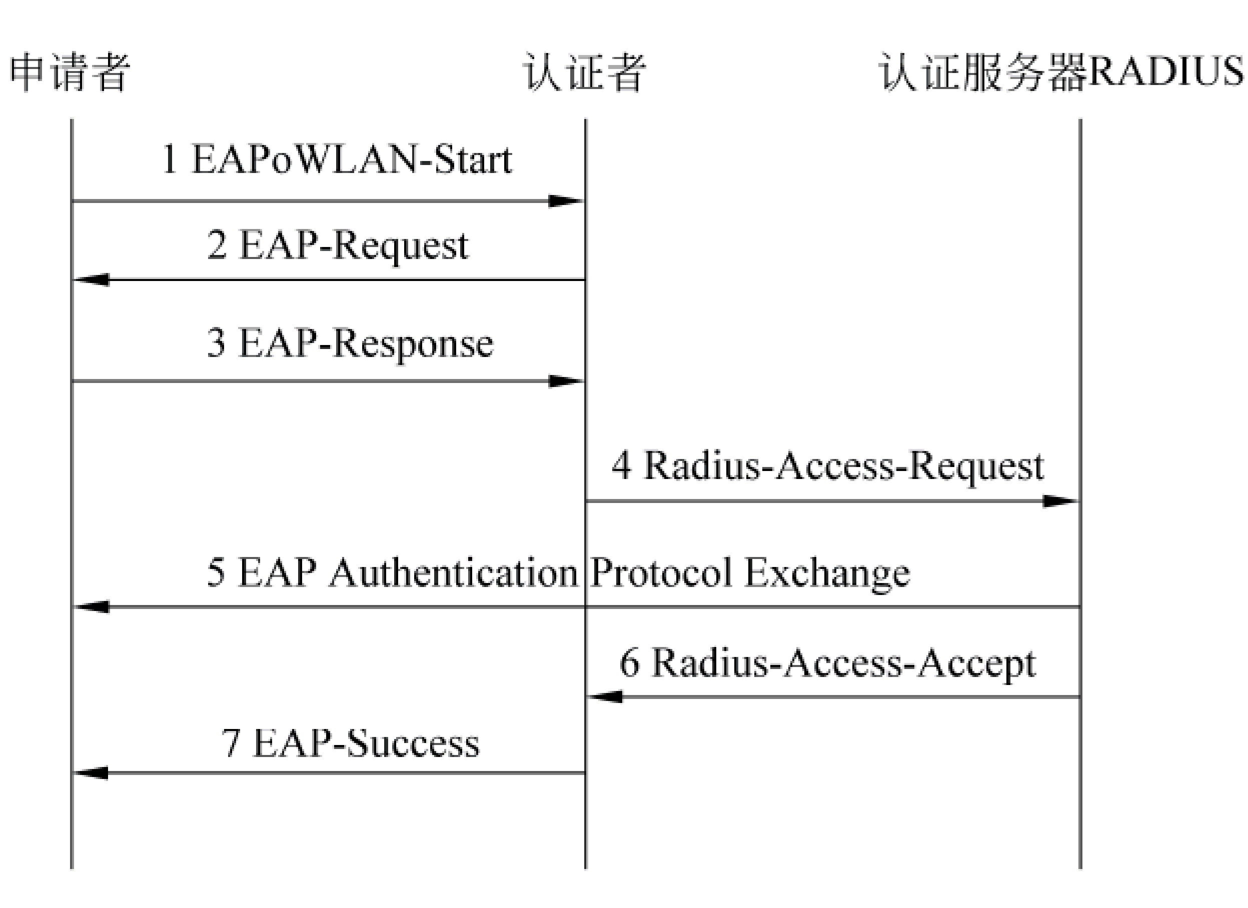


图 5.8　IEEE 802.1X 认证过程

(1) 请求者向认证者发送 EAP-Start 帧,启动认证流程。

(2) 认证者发出请求,要求请求者提供相关身份信息。

(3) 请求者回应认证者的请求,将自己的相关身份信息发送给认证者。

(4) 认证者将请求者的身份信息封装至 Radius-Access-Request 帧中,发送至 AS。

(5) RADIUS 服务器验证请求者身份的合法性,在此期间可能需要多次通过认证者与请求者进行信息交互。

(6) RADIUS 服务器告知认证者认证结果。

(7) 认证者向请求者发送认证结果,如果认证通过,那么认证者将为请求者打开一个受控端口,允许请求者访问认证者所提供的服务,反之,则拒绝请求者的访问。

扩展认证协议 EAP 是一种封装协议,在具体应用中可以选择 EAP-TLS、EAP-MD5、EAP-SIM、EAP-TTLS、EAP-AKA 等任何一种认证协议。不同的具体认证协议具有不同的安全性。其中,EAP-TLS 以数字证书作为凭证相互认证,是 EAP 种类中唯一基于非对称密码的认证方式。EAP-TLS 的消息交换可以提供远程 VPN 客户端和验证程序之间的相互身份验证、加密方法的协商和加密密钥的确定,提供了最强大的身份验证和密

钥确定方法。EAP-SIM 以移动电话的 SIM 卡(用户识别模块卡)进行身份验证。

下面就一个具体的 EAP-MD5 给出一个实例来说明 802.1X 协议,如图 5.9 所示。EAP-MD5 使用与基于 PPP 的 CHAP 相同的挑战/应答协议,但是挑战和应答是作为 EAP 消息发送的。EAP-MD5 是一种单向认证机制,不支持加密密钥的生成。EAP-MD5 的典型用法是通过使用用户名和密码对远程 VPN 客户端的凭据进行身份验证。

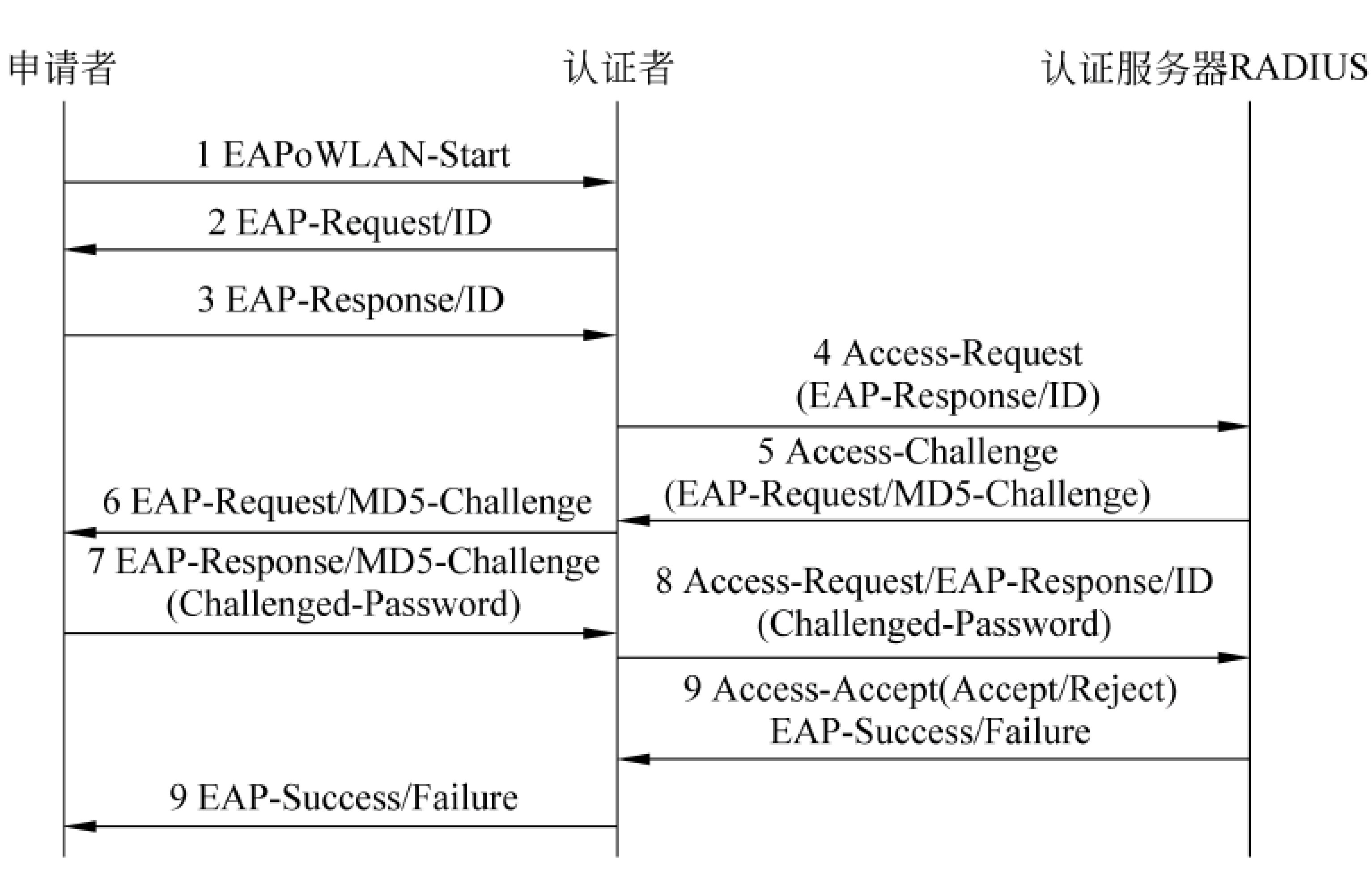


图 5.9　基于 EAP-MD5 的 802.1X 认证流程

认证流程如下:

(1) 请求者向认证者发送一个 EAPoWLAN-Start 报文,开始 802.1X 认证接入;

(2) 认证者向请求者发送 EAP-Request/ID 报文,要求请求者将用户名传来;

(3) 请求者响应一个 EAP-Response/ID 给认证者,其中包括用户名;

(4) 认证者将 EAP-Response/ID 报文封装到 RADIUS Access-Request 报文中,发送给认证服务器;

(5) 认证服务器产生一个挑战 Challenge,将 RADIUS Access-Challenge 报文发送给认证者,其中包含有 EAP-Request/MD5-Challenge;

(6) 认证者将 EAP-Request/MD5-Challenge 发送给请求者,要求认证请求者;

(7) 请求者收到 EAP-Request/MD5-Challenge 报文后,将密码和 Challenge 做 MD5 算法后的 Challenged-Password,EAP-Response/MD5-Challenge 应答给认证者;

(8) 认证者将 Challenge,Challenge-Password 和用户名一起送到 RADIUS 服务器,RADIUS 服务器进行认证;

(9) RADIUS 服务器根据用户信息,做 MD5 算法,判断用户是否合法,然后应答认证成功/失败报文到认证者。

3. IEEE 802.1X 认证的优点

IEEE 802.1X 协议能适应现代(无线)网络用户数量急剧增加和业务多样性的要求,具有以下优点:

(1) 协议实现简单。IEEE 802.1X 协议为两层协议,不需要到达第三层,因而对设备的整体性能要求不高,可有效降低建网成本。不需要进行协议间的多层封装,去除了不必

要的开销和冗余。采用 802.1X 方式,用户可用有线网络的速度进行工作,一台服务器能够在多个接入点之间处理多达 20000 个用户的认证。同时,网络综合造价成本低,保留了传统 AAA 认证的网络架构,可以利用现有的 RADIUS 设备。

(2) 业务灵活。IEEE 802.1X 的认证体系结构中采用了"受控端口"和"非受控端口"的逻辑功能,用户通过认证后,业务流和认证流实现分离,通过认证后的数据包是不需封装的纯数据包,通过受控端口进行交换,因而业务可以很灵活(尤其在开展宽带组播等业务时有很大的优势),易于支持多业务和新兴流媒体业务。

(3) 安全可靠。具体表现在如下几个方面。

① 用户身份识别取决于用户名、口令、数字证书等,而不是 MAC 地址,从而可实现基于用户的认证、授权和计费。

② 支持可扩展的认证、非口令认证,如公钥证书和智能卡、互联网密钥交换协议(IKE)、生物测定学、信用卡等,同时也支持口令认证,如一次性口令认证、通用安全服务应用编程接口方法(包括 Kerberos 协议)。

③ 协议的有些版本支持双向认证,可有效防止了中间人攻击和假冒接入点 AP,还可防范地址欺骗攻击、目标识别和拒绝服务攻击等,并支持针对每个数据包的认证(完整性保护)。并可以在不改变网络接口卡的情况下,插入新的认证(以及密钥管理)方法。

④ 与 PPPoE 和 Web/Portal 认证方式相比,消除了网络瓶颈,减轻了网络封装开销,降低了建网成本。PPPoE 认证中,认证系统必须将每个包进行拆解才能判断和识别用户是否合法,一旦用户增多或数据包增大,封装速度成为网络瓶颈,大量的拆包和封包过程导致设备昂贵。Web/Portal 认证是基于业务类型的认证,需要安装浏览器才能完成,且是应用层认证,且认证连接性差,不容易检测用户离线,基于时间的计费难以实现。

5.2.3 IEEE 802.11i 接入协议

针对 IEEE 802.11 WEP 安全机制所暴露出的安全隐患,IEEE 802 工作组于 2004 年初发布了新一代安全标准 IEEE 802.11i(也称为 WPA2,Wi-Fi Protected Access,以及 RSN,Robust Security Network)[8]。

首先该协议将 IEEE 802.1X 协议引入到 WLAN 安全机制中,增强了 WLAN 中身份认证和接入控制的能力;其次,增加了密钥管理机制,可以实现密钥的导出及密钥的动态协商和更新等,大大增强了安全性。IEEE 802.11i 提出了两种加密机制:TKIP 协议(Temporal Key Integrity Protocol)和 CCMP 协议(Counter Mode/CBC-MAC Protocol)。TKIP 是一种临时过渡性的可选方案,兼容 WEP 设备,可在不更新硬件设备的情况下升级至 IEEE 802.11i;而 CCMP 机制则完全废除了 WEP,采用加密算法 AES(Advanced Encryption Standard)来保障数据的安全传输,但是 AES 对硬件要求较高,CCMP 无法在现有设备的基础上通过直接升级来实现(需要更换硬件设备),它是 IEEE 802.11i 机制中要求必须实现的安全机制,是 802.11i 的关键技术。

另外,在 802.11i 制定的 TKIP 作为过渡期间,对迫切需要解决安全问题的商家和用户而言,标准的批准滞后是无法容忍的,于是 Wi-Fi 联盟推出了 WPA,WPA 不是一个正式的标准,只是过渡到 802.11i 的中间标准。

下面在本节介绍 IEEE 802.11i 接入机制，下一节将介绍 IEEE 802.11i 加密机制 TKIP 和 CCMP。

1. 接入流程

IEEE 802.11i 协议接入流程一般包括发现、认证和密钥协商 3 个阶段，其中每个阶段又由若干子步骤组成，共同实现 IEEE 802.11i 协议功能，如图 5.10 所示，图中 STA 表示工作站，AP 表示接入点，AS 是认证服务器，具体流程如下：

(1) 发现阶段。STA 启动后，通过被动侦听 AP 发送的信标帧，或主动发出探寻请求来检测周围是否有可以接入的 AP 并获取相关安全参数。若检测到多个可选的 AP，就选其中一个，与该 AP 进行认证和关联。该阶段的认证方式包括两种，开放认证和共享密钥认证，共享密钥认证为可选认证方法。该阶段的认证不可靠，需要在后续过程中强化。

(2) 认证阶段。IEEE 802.11i 协议引入 IEEE 802.1X 协议进行认证，目的是在发现阶段构建的关联和不可靠认证的基础上，利用 IEEE 802.1X 协议强化身份认证，确保对网络资源的访问是合法的。EAP-TLS 是一种双向认证机制，也是目前 802.11i 的默认认证协议。同时在该阶段，在 STA 和 AS 间生成成对主密钥(Pairwise Master Key，PMK)，PMK 为 IEEE 802.11i 协议密钥建立体系的基础，PMK 从 AS 安全传递至 AP。

(3) 密钥协商阶段。密钥协商阶段包括进行单播密钥协商的 4 步握手协议和进行组播密钥握手协议。该阶段的目的是，在生成 PMK 的基础上，导出单播密钥和组播密钥保护后续数据的安全传输。

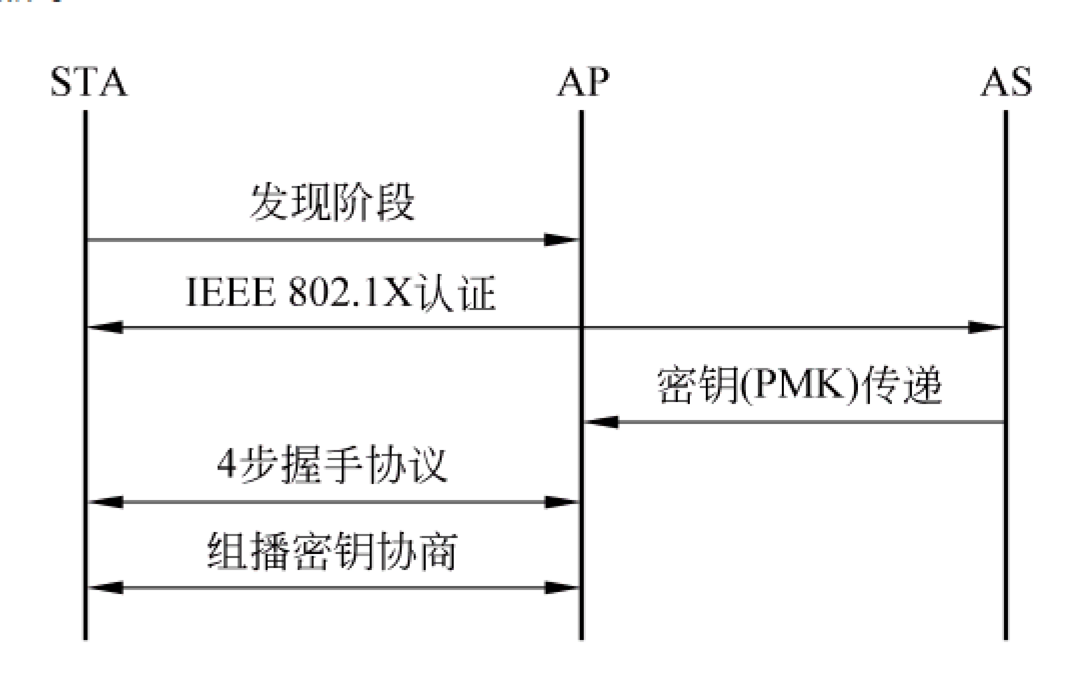


图 5.10　IEEE 802.11i 接入流程(接入认证、密钥传递、密钥协商)

2. 密钥协商协议与密钥管理

早期的 EAP 消息交换使得在 STA 和 AP 之间建立了 PMK，所谓成对(Pairwise)密钥，是因为它在 STA 和 AP 间共享；所谓主(Master)密钥，是因为其不直接用于消息加密或完整性保护，而是从 PMK 生成加密密钥和完整性密钥。更确切地说，STA 和 AP 均从 PMK 导出 4 个密钥：数据加密密钥 TK(Temporal Key，16 字节)、数据完整性密钥 MIC Key(AP 和 STA 各 8 字节)、密钥加密密钥 KEK(16 字节)、密钥完整性密钥 KCK(16 字节)，这 4 个密钥一起称为 PTK(Pairwise Transient Key，成对临时密钥)。(AES-CCMP 中利用相同的密钥进行数据加密和完整性保护，因此在 AES-CCMP 中，PTK 仅由 3 个密钥组成)。此外，从 PMK 导出的 PTK 与 AP 和 STA 的 MAC 地址以及双方产生的随机数 Nonce 有关。PTK 由密码学安全的 Hash 函数产生，其输入参数是 AP Nonce (ANonce)，STA Nonce (SNonce)、AP MAC 地址以及 STA MAC 地址的连接

(Concatenation)。

STA 和 AP 交换各自随机数使用的协议称为 4 路握手协议(Four-way Handshaking Protocol)。此协议向对方证明自己拥有 PMK,并生成 PTK。4 路握手协议的描述如下:

(1) AP 发送其随机数 ANonce 给 STA。当 STA 收到 ANonce 后,可计算出 PTK。

(2) STA 发送其随机数 SNonce 给 AP。此消息携带一个消息完整码 MIC,由 STA 使用刚刚计算的 PTK 中的密钥完整性密钥计算而来。接收该随机数后,AP 可计算出 PTK。因此,AP 可利用计算出的 PTK 中的密钥完整性密钥验证 MIC。如果认证成功,则 AP 相信 STA 拥有 PMK。

(3) AP 发送一个包含 MIC 的消息给 STA。MIC 由 PTK 的密钥完整性密钥计算得来。如果 STA 验证 MIC 通过,则其相信 AP 也拥有 PMK。该消息包含序列号以检测重放攻击。此消息告知 STA,AP 已经准备好加密所有数据包的密钥。

(4) STA 确认接收到第三个消息。该确认也意味着 STA 准备好加密所有数据包。

一旦得到 PTK,则 STA 和 AP 之间的数据包将得到数据加密密钥和数据完整性密钥的保护。然而,这些密钥不能用于保护由 AP 发送的广播消息。保护广播消息的密钥必须被"所有"STA 和 AP 已知,因此 AP 产生额外的组密钥称为 GTK(Group Temporal Key,组播临时密钥)。GTK 包含一个组播加密密钥和组播完整性密钥,并且将用其给 STA 的密钥加密密钥加密,然后分别发送给每一个 STA。

容易看到,IEEE 802.11i 接入协议是典型的认证密钥协商 AKA 协议。另外,由成对主密钥 PMK 导出后续的会话加密密钥和数据完整性密钥的方法是一种典型的密钥分层管理的方法,密钥的分层管理可提高密钥的安全性(抗密钥泄漏的健壮性),在安全设计中很常见。

*5.2.4 IEEE 802.11i TKIP 和 CCMP 协议

1. TKIP 加密机制

TKIP 协议是 IEEE 802.11i 标准采用的过渡安全解决方案,它可以在不更新硬件设备的情况下,通过软件升级实现安全性的提升。TKIP 与 WEP 一样都是基于 RC4 加密算法,但是为了增强安全性,初始化向量 IV 的长度由 24 位增加到 48 位,并称之为 TSC(TKIP Sequence Counter),同时对 WEP 协议进行了改进,新引入了 4 种机制来提升安全性:

(1) 防止出现弱密钥的单包密钥(Per-Packet Key,PPK)生成算法。

(2) 使用 Michael 算法防止数据遭非法篡改的消息完整性校验码(MIC)。

(3) 防止重放攻击的具有 48 位序列号功能的 IV(即 TSC)。

(4) 可生成新鲜的加密和完整性密钥,防止 IV 重用的再密钥(Rekeying)机制。

TKIP 的加密过程(如图 5.11 所示)包括以下几个步骤(需要熟悉 802.11MAC 帧结构):

(1) MAC 协议数据单元(Medium Access Control Protocol Data Unit,MPDU)的生成:首先发送方根据源地址(SA)、目的地址(DA)、优先级(Priority)和 MAC 服务数据单元(MAC Service Data Unit,MSDU),利用 MIC 密钥(MIC Key)通过 Michael 算法计算

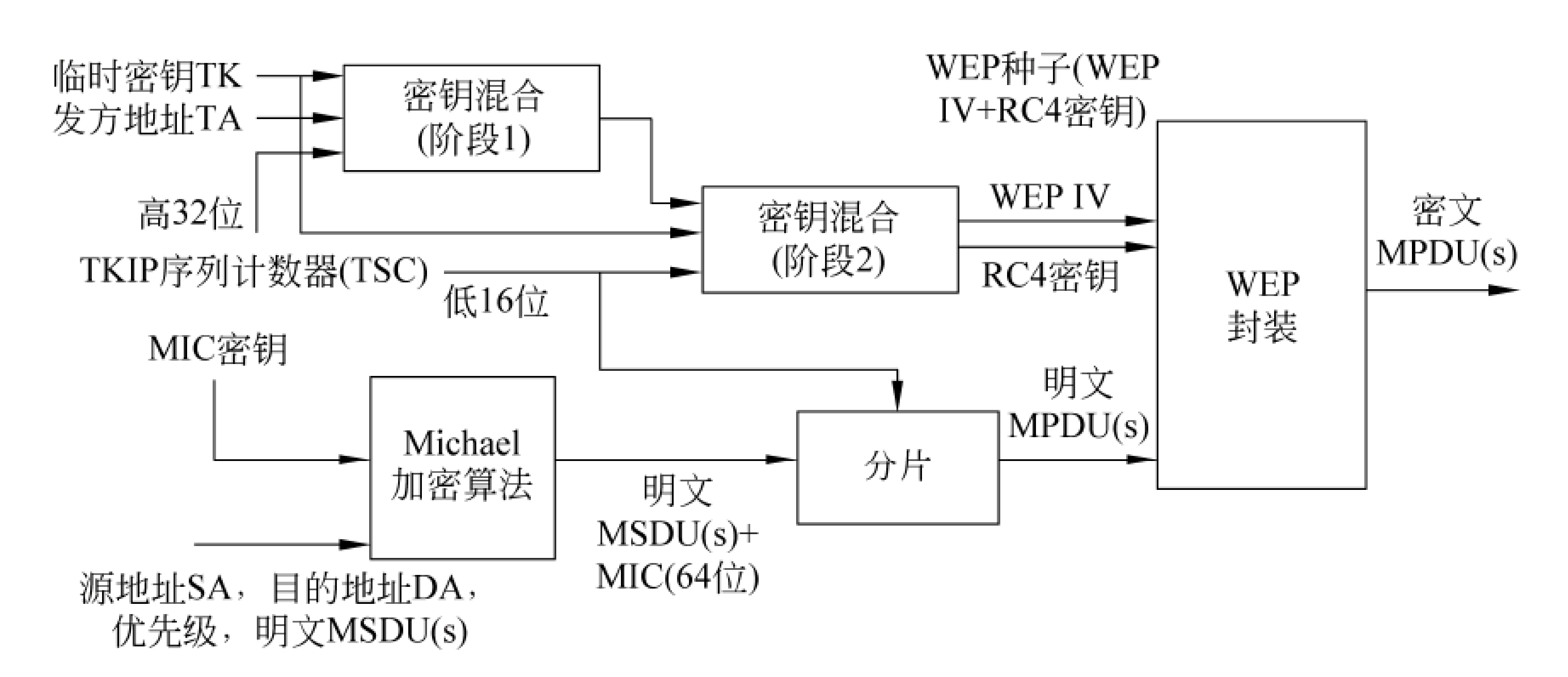


图 5.11　TKIP 加密过程

出消息完整性校验码(MIC)，并将 MIC 添加到 MSDU 后面，一起作为 WEP 算法的加密对象，如果 MSDU 加上 MIC 的长度超出 MAC 帧的最大长度，可以对 MPDU 进行分片。

(2) WEP 种子的生成：TKIP 将临时密钥(Temporal Key，TK)、发方地址(TA)及 TKIP 序列计数器(TSC)经过两级密钥混合(Key Mixing)函数(一种 AES 中导出的 S 盒，具有非线性和运算快的特点)后，得到用于 WEP 加密的 WEP 种子(WEP Seeds)。对于每个 MPDU，TKIP 都将计算出相应的 WEP 种子(同一 MPDU 的分片使用同一个 WEP 种子)。

(3) WEP 封装(WEP Encapsulation)：TKIP 计算得出的 WEP 种子分解成 WEP IV 和 RC4 密钥的形式，然后把它和对应的 MPDU 一起送入 WEP 封装器(一种硬件加密模块)进行加密，得到密文 MPDU 并按规定格式封装后发送。

2. CCMP 加密机制

由于序列密码 RC4 算法并不安全，于是考虑采用分组密码算法。AES 是美国 NIST 制定的用于取代 DES 的分组加密算法，CCMP 是基于 AES 的 CCM 模式(Counter Mode/CBC-MAC Mode)，完全取代了原有 WEP 加密，能够解决 WEP 加密中的不足，可以为 WLAN 提供更好的加密、认证、完整性和抗重放攻击的能力，是 IEEE 802.11i 中强制要求实现的加密方式，同时也是 IEEE 针对 WLAN 安全的长远解决方案。CCMP 加密过程如图 5.12 所示。

(1) 为保证每个 MPDU 都具有新鲜的包号码(Packet Number，PN)，增加 PN 值，使得每个 MPDU 对应一个新的 PN，这样即使对于同样的临时密钥，也不会出现相同的 PN。

(2) 用 MPDU 帧头的各字段为 CCM 生成附加鉴别数据(Additional Authentication Data，AAD)，CCM 为 AAD 的字段提供完整性保护。

(3) 用 PN、A2 和 MPDU 的优先级字段计算出 CCM 的使用一次的随机数(Nonce)。其中 A2 表示地址 2，优先级字段作为保留值设为 0。

(4) 用 PN 和 Key Id 构建 8 字节的 CCMP 头。

(5) 由 TK、AAD、Nonce 和 MPDU 数据生成密文，并计算 MIC 值。最终的消息由 MAC 头、CCMP 头、加密数据以及 MIC 连接而成。

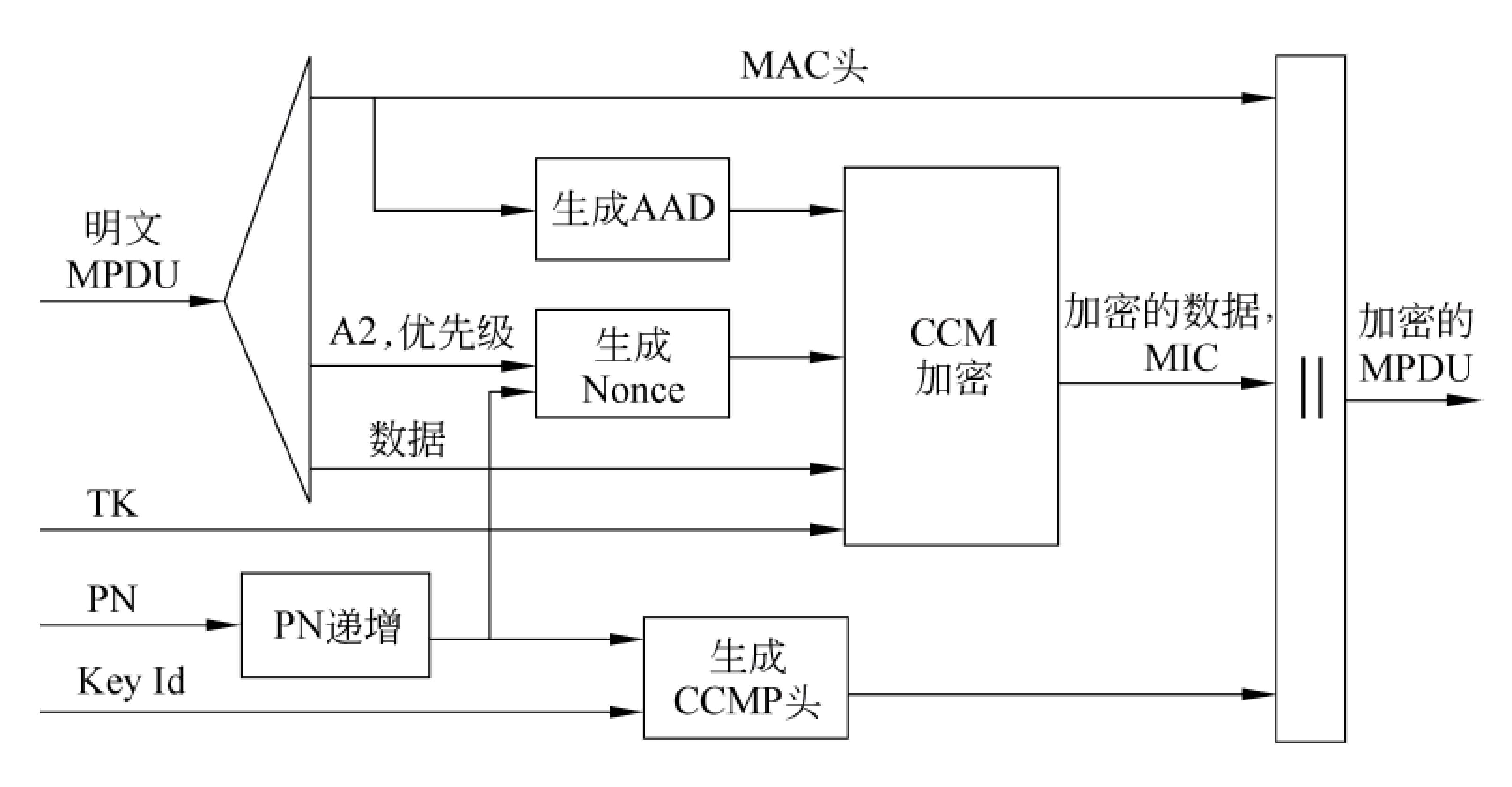


图 5.12 CCMP 加密过程

3. WEP、TKIP 和 AES-CCMP 的比较

TKIP 与 AES-CCMP 都是用数据加密和数据完整性密钥保护 STA 和 AP 之间传输数据包的完整性和保密性。然而,它们使用了不同的密码学加密算法。TKIP 与 WEP 一样使用 RC4,但是与 WEP 不同的是,提供了更多的安全性。TKIP 的优势为通过固件升级,可在旧 WEP 硬件上运行。AES-CCMP 使用 AES 算法,需要支持 AES 算法的新硬件,但与 TKIP 相比,提供了一个更清晰、更健壮的解决方案。TKIP 修复 WEP 中的缺陷包括如下。

(1) 完整性:TKIP 引进了一种新的完整性保护机制,使用 Michael 算法。Michael 运行在服务数据单元(SDU)层,可在设备驱动程序中实现。

(2) 检测重放攻击:TKIP 使用新 IV 机制(TSC)作为序列号。TSC 初始化后,每发送一个消息后自增。接收者记录最近接收消息的 TSC。如果最新接收消息的 TSC 值小于存储的最小 TSC 值,则接收者扔掉此消息;如果 TSC 大于存储的最大 TSC 值,则保留此消息,并且更新其存储的 TSC 值,如果刚收到消息的 TSC 值介于最大值和最小值之间,则接收者检查 TSC 是否已经存储;如果有记录,则扔掉此消息;否则,保留此消息,并且存储新的 TSC。

(3) 保密性:WEP 加密的主要问题为 IV 空间太小,并且没有考虑 RC4 中存在的弱密钥。为了克服第一个问题,在 TKIP 中,IV 从 24 位增加至 48 位。由于 WEP 硬件仍然期望一个 128 位的 WEP 种子(单包加密密钥)。因此 48 位 IV 与 128 位 TK 混合完后必须用某种方式压缩为 128 位。对弱密钥问题,在 TKIP 中单包加密密钥都不相同(因为 RC4 密钥不同,以及 IV 低 16 位不同)。因此,攻击者不能观察到具有使用相同密钥的足够数量的消息。

TKIP 的新 IV 机制及单包加密密钥(WEP 种子)的生成如图 5.13 所示。48 位 IV(即 TSC)分为高 32 位和低 16 位。IV 的高 32 位与 128 位临时密钥和 STA 的 MAC 地址相混合(密钥混合阶段 1)。然后,将此计算结果与 IV 的低 16 位相混合(密钥混合阶段 2),得到 104 位 RC4 密钥。TKIP 的 WEP 种子由 RC4 密钥、IV 的低 16 位(分成两个字节)及一个虚假填充字节 d 拼接而来。

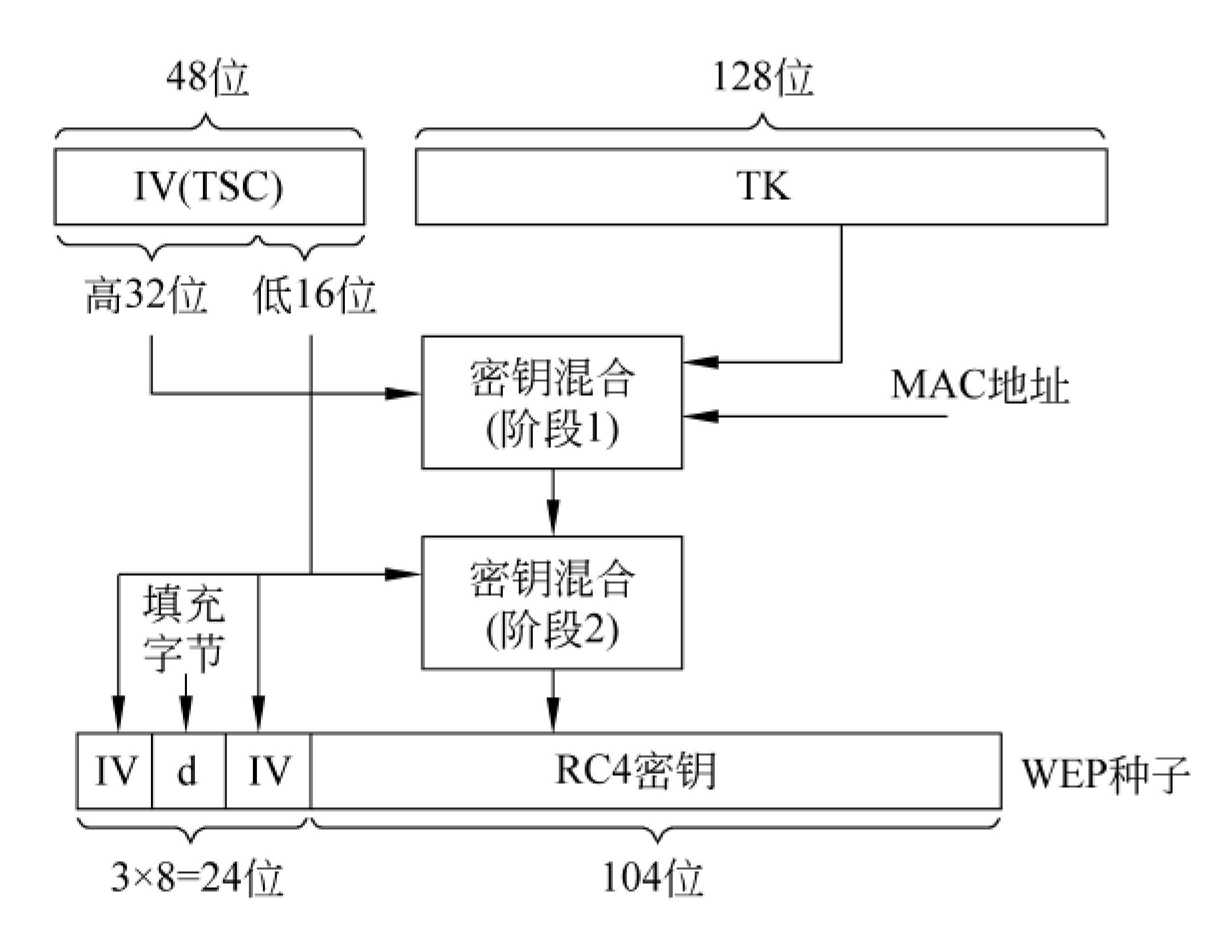


图 5.13　TKIP 中生成 WEP 种子(WEP IV+RC4 密钥)

因为不必为兼容 WEP 硬件所束缚,AES-CCMP 的设计要比 TKIP 简单。它放弃 RC4,使用 AES 分组加密,并定义了一个新工作模式,称为 CCM。CCM 由两种工作模式结合而来:CTR 加密模式和 CBC MAC 模式。在 CCM 模式中,消息发送方计算出消息的 CBC MAC 值,并将其附加到消息上面,然后将其用 CTR 模式加密。CCM 模式确保了保密性和完整性。重放攻击检测由消息的序列号来保证,通过将序列号加入到 CBC MAC 计算的初始块中来完成。

通过上述比较,可体会到安全工程设计中往往需要考虑非技术因素(如前向兼容性等经济成本因素)。同时利用 MAC 帧格式中的字段作为影响密钥生成的因素也是工程设计中的特点之一(充分利用应用环境的上下文作为安全设计的关联因素)。

5.2.5　WAPI 协议

针对 IEEE 802.11 WEP 安全机制的不足,2003 年我国也提出了一个无线局域网安全标准(Wireless LAN Authentication and Privacy Infrastructure,WAPI,无线局域网认证和保密基础结构),这也是我国首个在无线网络通信领域自主创新并拥有知识产权的安全接入技术标准,是我国首个无线通信网络安全领域的国际标准(ISO/IECJTC1/SC6 会议上通过),具有重要的历史意义和战略影响。

WAPI 由 WAI(WLAN Authentication Infrastructure)认证基础结构和 WPI(WLAN Privacy Infrastructure)隐私基础结构两部分组成,WAI 和 WPI 分别实现对用户身份的鉴别和对传输的数据加密。

WAI 认证结构其实类似于 IEEE 802.1X 结构,也是基于端口的认证模型。采用公开密钥密码体制,利用数字证书(独立设计的数据结构、不兼容 X.509 证书格式)来对 WLAN 系统中的 STA 和 AP 进行认证。WAI 定义了一种名为 ASU(Authentication Service Unit,认证服务单元)的实体,用于管理参与信息交换各方所需要的证书(包括证书的产生、颁发、吊销和更新),相当于 PKI 中的 CA 的角色。通常 ASU 的物理形态为认证服务器(Authentication Server),AS 逻辑上包含了 ASU 的功能。证书里包含有证书

持有者的标识、公钥和证书颁发者(ASU)的签名(这里的签名采用的是国家商用密码管理办公室颁布的椭圆曲线数字签名算法),证书是网络设备的数字身份凭证。

整个系统由移动终端(STA)、接入点(AP)和认证服务器(AS)组成,其中AS含有ASU可信第三方,用于管理消息交换中所需要的数字证书。AP提供STA连接到AS的端口(即非受控端口),确保只有通过认证的STA才能使用AP提供的数据端口(即受控端口)访问网络。

WAPI整个过程由证书鉴别、单播密钥协商和组播密钥通告(合称密钥协商阶段)3部分组成,如图5.14所示。

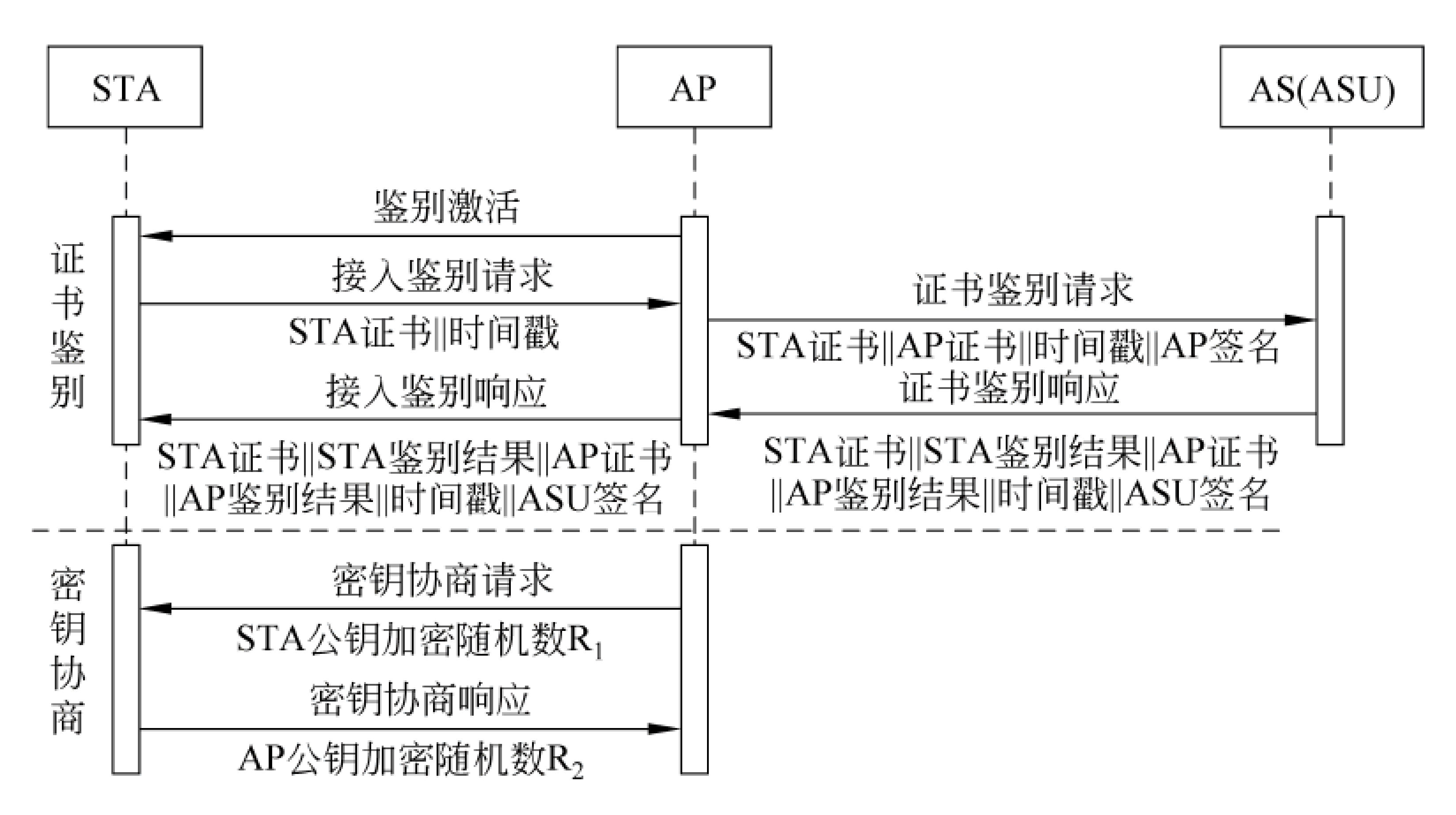


图5.14　WAPI的接入认证与密钥协商过程

证书鉴别阶段中,STA、AP提交各自证书给AS,AS验证它们的有效性后返回鉴别响应。STA和AP验证AS对响应消息的数字签名,获得验证结果,并认证AS的合法性。在WAI协议中,STA、AP无需下载证书列表或在线验证证书状态,由AS统一进行证书有效性验证,同时AS担当STA、AP等实体证书发放、撤销和管理,这种简化的集中化管理,无需额外的权威授权中心CA,架构设计非常简单。

下面介绍具体的过程,为了简便,从下面开始AS和ASU是等同对待的。

1. 证书鉴别过程

(1) 鉴别激活。当STA关联至AP时,由AP向STA发送鉴别激活以启动整个鉴别过程。

(2) 接入鉴别请求。STA向AP发出接入鉴别请求,即将STA证书与STA当前系统时间一同发送给AP。

(3) 证书鉴别请求。AP收到STA接入鉴别请求后,首先记录鉴别请求时间,然后向AS发出证书鉴别请求,即将STA证书、接入鉴别请求时间、AP证书,以及AP的私钥对它们的签名,组成证书鉴别请求发送给AS。

(4) 证书鉴别响应。AS收到AP的证书鉴别请求后,验证AP的签名和AP证书的有效性,若不正确,则鉴别过程失败,否则进一步验证STA证书,验证完毕后,AS将STA证书和STA证书鉴别结果、AP证书和AP证书鉴别结果,以及ASU对它们的签名,组

成证书鉴别响应发回给 AP。

(5) 接入鉴别响应。AP 验证 AS 返回的证书鉴别响应的签名,得到 STA 证书的鉴别结果,根据此结果对 STA 进行接入控制,从而完成了对 STA 的认证。AP 将收到的证书鉴别响应回送至 STA。STA 验证 ASU 的签名后,得到 AP 证书的鉴别结果,根据该鉴别结果决定是否接入该 AP,从而完成对 AP 的认证。

至此 STA 与 AP 之间完成了认证过程。容易看到,该认证方法是基于公钥密码学在可信第三方存在条件下的双向认证协议。也就是说,WAPI 中 STA 和 AP 的双向认证其实是指通过可信第三方 AS 的认证:当 STA 关联 AP 时,AP 和 STA 的证书都要被 AS 来鉴别。只有鉴别成功,AP 才允许 STA 接入,同时 STA 也才允许通过该 AP 收发数据。这种认证对于采用“假”AP 的攻击方式具有很强的抵御能力。这种认证其实也实现了实现 AS 与 STA、AS 与 AP 的双向认证(严格意义上讲,WAPI 中的这种认证只是对证书真实性的鉴别,而不是对证书所有者的认证。WAPI 标准没有称为实体认证是恰当的,其实 WAPI 并没有完全实现对 STA 的认证)。认证完成后,AP 向 STA 发送密钥协商请求分组开始与 STA 协商单播密钥。

WAPI 在认证方面(即 WAI)具有以下几个重要特点:具有自主知识产权;完整的 STA 和 AP 双向认证;协议交换消息少,通信效率高;集中式或分布集中式认证管理;灵活多样的证书管理与分发体制;可支持多证书,方便用户多处使用,充分保证其漫游功能;认证服务单元易扩充,支持用户的异地接入。

WAI 仍然存在着一些不完善之处。例如 STA 将自己的证书以明文形式发送给 AP 及 AS,因而会暴露用户的身份信息,无法实现认证时的匿名性。另外,AP 和 AS 都易成为计算瓶颈及易遭受拒绝服务。尤其需要注意的是,实际上有可能没有认证 STA 本身(只鉴别了 STA 证书的合法性)。下面简要说明一下。

因为 AS 验证了 AP 的签名,因此 AP 的合法性可保证。AP 和 STA 也验证了 AS 的签名,AS 的合法性可保证。但是 AS 只验证了 STA 的证书合法性,因此,STA 的合法性需要等到密钥协商完成后,看 STA 是否能够正确使用自己的私钥解密 AP 的随机数,即是否具有欲协商的共享密钥才能确定。通过判断 STA 是否能正确生成合法会话密钥的对 STA 的认证方法,有时称之为“隐性认证”。这个在下面介绍了单播密钥协商过程就清楚了。

2. 单播密钥协商过程

(1) 密钥协商请求。AP 采用伪随机数生成算法生成伪随机数 R_1,利用 STA 的公钥将其进行加密。AP 将密钥协商标识、单播密钥索引、加密信息和安全参数索引等用自己的私钥生成签名后发送给 STA。

(2) 密钥协商响应。STA 检查当前状态、安全参数索引和 AP 签名的有效性。然后查看分组是证书认证成功后的首次密钥协商还是密钥更新协商请求,并且相应地对密钥协商标识字段值进行比较。然后用自己的私钥解密得到 R_1,STA 产生 R_2,将 $R_1 \oplus R_2$ 进行扩展得到单播会话密钥。STA 将单播密钥索引、下次密钥协商标识、消息鉴别码、用 AP 公钥加密的 R_2 等发送给 AP。

AP 收到密钥协商响应消息后,使用私钥解密得到 R_2,扩展 $R_1 \oplus R_2$ 得到单播会话密

钥和消息鉴别(数据完整性)密钥,计算消息鉴别码,将其和响应分组中的消息鉴别码字段进行比较。最后比较会话算法标识,判断下次密钥协商标识是否单调递增,保存下次密钥协商表示作为下次单播密钥更新时的密钥协商标识。

在单播密钥协商完成后,开始组播密钥协商过程。其过程类似,具体描述可参见相关文献。

下面简要比较一下 802.11i EAP-TLS 与 WAPI 的不同。

EAP-TLS 认证是基于 STA 和 RADIUS 服务器的双向认证,并且使用了数字证书,但并没有对选用的 AP 进行充分的验证,这会带来一定的安全隐患;同时在 EAP-TLS 认证的最后由 AP 发送给 STA 的报文"EAP-Success"采用明文传送,容易被攻击者利用,从而达到欺骗客户端,进而导致引起遭到拒绝服务攻击和中间人攻击。

WAPI 中也使用了数字证书(虽然格式不同),AP 和 AS 间相互认证,AS 和 STA 间相互认证,AP 和 AS 的合法性可保证,但 STA 只验证了证书,STA 的身份是通过密钥协商"隐性认证"的。

WAPI 的设计思路与 802.11i 截然不同,802.11i 为了兼容性,组合了一些现有的有线网络安全协议,而 WAPI 重新设计了 WLAN 安全结构,因此其认证协议消息交换轮数与 802.11i 中的 EAP-TLS 认证协议相比要少得多;而且 WAPI 的密钥协商协议也非常简单,只需要进行两轮交换消息,不像 802.11i 中是四次握手。

通过比较可以发现,这某种程度上体现了实际网络安全工程设计中除了安全性外,还需要考虑更多的因素,如历史兼容性、投资和成本等非安全因素。另外也可以看出,密码学安全协议仅仅提供的是协议的最简单形式及其安全保障,在具体实际的应用中形态可能是多种多样的。

*5.2.6 SMS4 对称密码算法

特别值得注意的是,WAPI 中使用的加密算法是我国自己制定的分组加密算法 SMS4。2006 年我国国家密码管理局公布了 WAPI 中使用的 SMS4 密码算法[9],该算法是我国拥有自主知识产权的加密算法。这是我国第一次公布自己的商用密码算法,意义重大,标志着我国商用密码管理更加科学化和与国际接轨。

另外,国家密码管理局公告(第 7 号,2006 年 1 月 6 日)[10]中要求:无线局域网产品须采用 SMS4 作为对称密码算法。鉴于上述特殊性,下面给予简介。

1. 基本参数与运算

SMS4 分组长度为 128 位,密钥长度为 128 位,加密和密钥扩展算法都采用 32 轮迭代结构。它以字节和字(32 位)为单位进行数据处理。

SMS4 中的基本运算为模 2 加和左循环移位。分别用$\oplus$和$<<<$表示。

2. SMS4 中的基本加密元素

(1) S 盒。S 盒的输入和输出都为一个字节,其本质是 8 位非线性置换,起混淆(Confusion)的作用。S 盒的设计(输入输出对照表)是公开的。假设输入为字节 x,输出为字节 y,S 盒的运算可表示为 y=S(x)。例如 S(00)=D6,S(01)=90 等(这里均为十六进制表示),S 盒的设计可参阅标准文档。

(2) 非线性变换 t。以字为单位,有 4 个 S 盒构成,实质上是 S 盒的并行计算。例如,设输入字为 $X=(x_0,x_1,x_2,x_3)$,输出字为 $Y=(y_0,y_1,y_2,y_3)$,则 $Y=t(X)=(S(x_1),S(x_2),S(x_3),S(x_4))$。

(3) 线性变换 L。以字为单位,主要起扩散(Diffusion)作用。设 L 的输入为字 X,输出为字 Y,则

$$Y=L(X)=X\oplus(X<<<2)\oplus(X<<<10)\oplus(X<<<18)\oplus(X<<<24)$$

(4) 合成变换 T。以字为单位。由非线性变换 t 和线性变换 L 复合而成。设输入为 X,输出为 Y,则有 $Y=T(X)=L(t(X))$。容易看到,合成变换同时起到混淆和扩散的作用。

3. 轮函数的设计

SMS4 的轮函数(Round Function)以字为处理单位。设轮函数 F 的输入为(X_0,X_1,X_2,X_3),共 4 个字,128 位,轮密钥 K 为 32 位(1 个字),如图 5.15 所示。

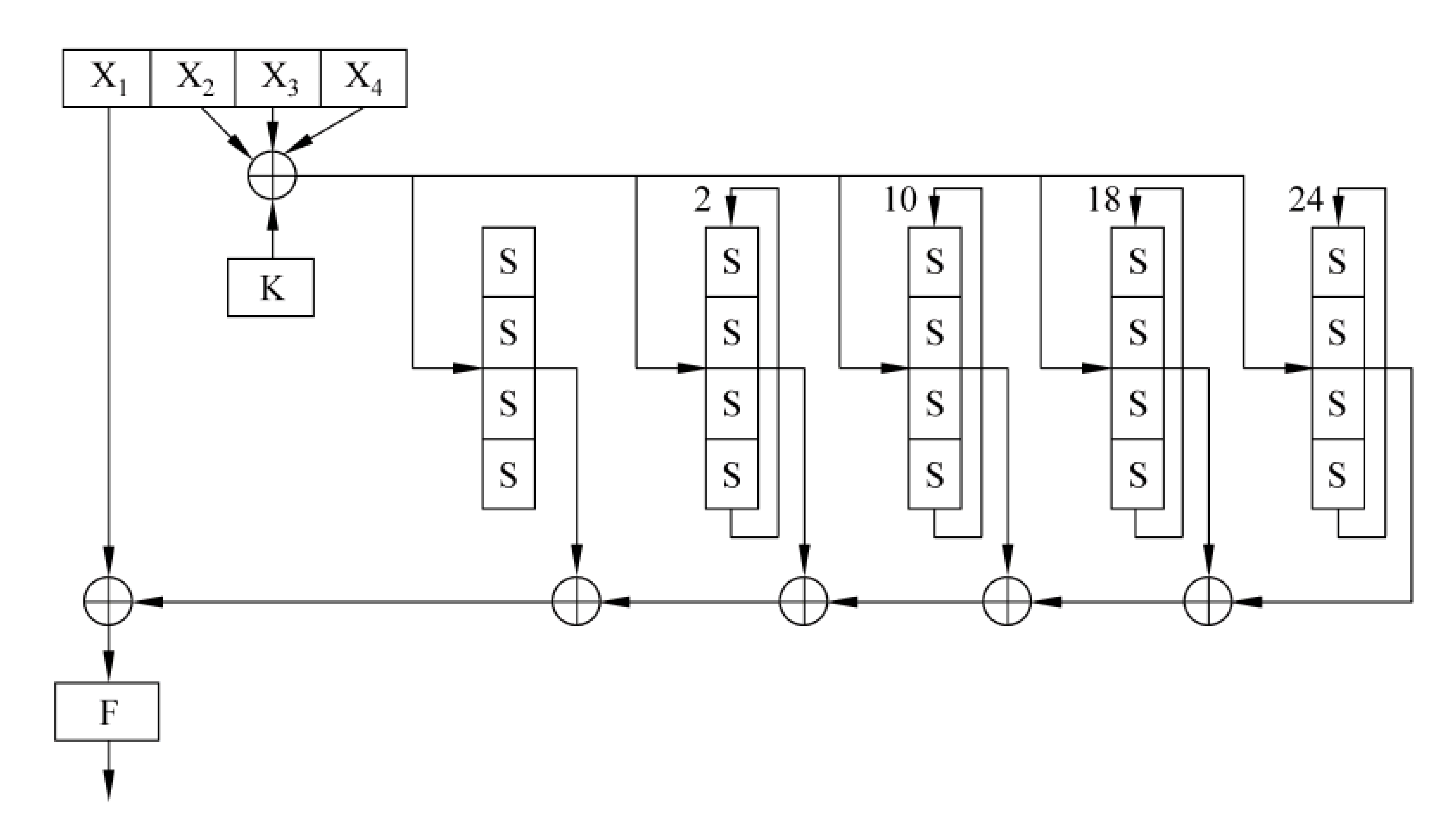


图 5.15　SMS4 的轮函数

即

$$\begin{aligned}&F(X_0,X_1,X_2,X_3,K)\\&=X_0\oplus T(X_1\oplus X_2\oplus X_3\oplus X_4\oplus K)=X_0\oplus L(t(X_1\oplus X_2\oplus X_3\oplus X_4\oplus K))\end{aligned}$$

若令 $B=X_1\oplus X_2\oplus X_3\oplus X_4\oplus K$,则

$$\begin{aligned}&F(X_0,X_1,X_2,X_3,K)\\&=X_0\oplus S(B)\oplus(S(B)<<<2)\oplus(S(B)<<<10)\oplus(S(B)\\&<<<18)\oplus(S(B))<<<24)\end{aligned}$$

4. 加密算法

共 32 轮加密,设输入明文为(X_0,X_1,X_2,X_3)4 个字,轮密钥为 K_i,$i=0,1,\cdots,31$,经过 32 轮运算后,输出密文(Y_0,Y_1,Y_2,Y_3),如图 5.16 所示。

即加密算法为

$$X_{i+4}=F(X_i, X_{i+1}, X_{i+2}, X_{i+3}, K_i)$$
$$=X_i \oplus T(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus K_i)$$
$$(Y_0, Y_1, Y_2, Y_3) = (X_{35}, X_{36}, X_{37}, X_{38})$$

这里的设计借用了密文反馈和流密码的思想。

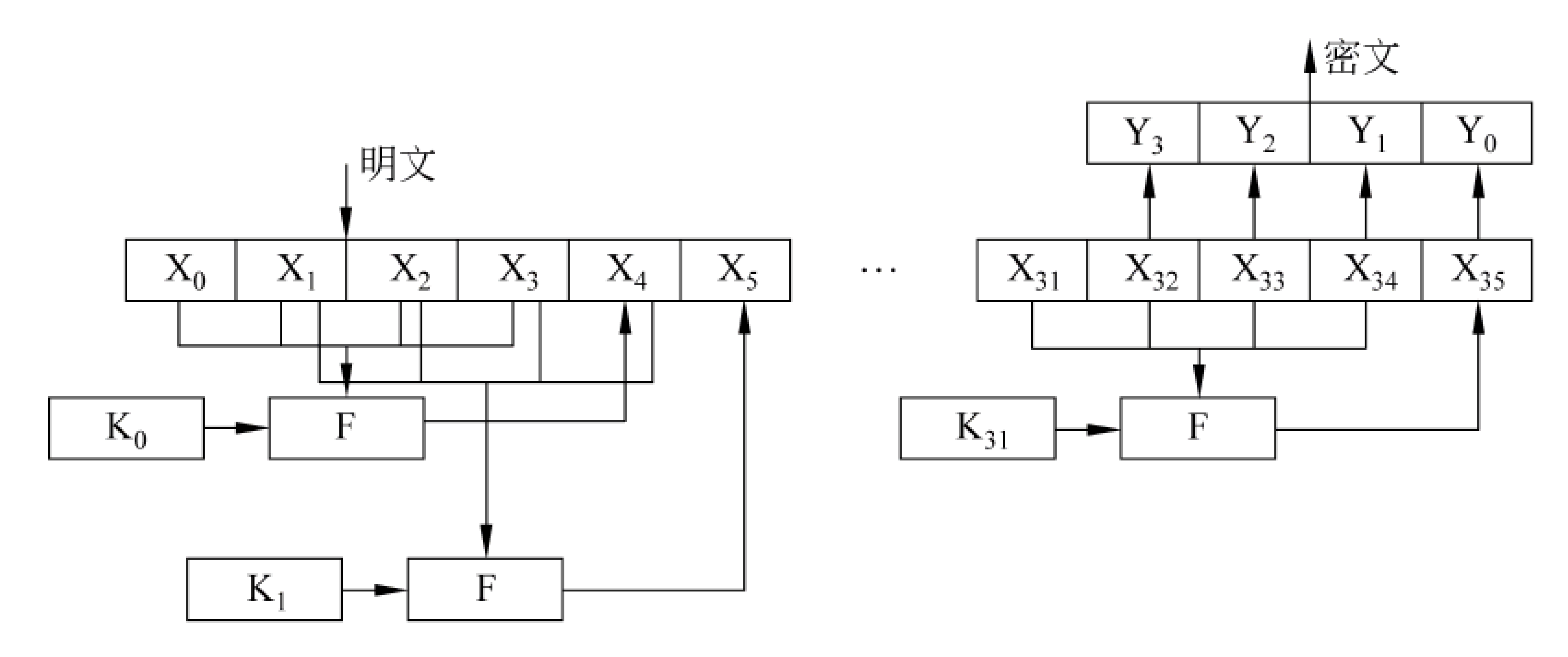


图 5.16　SMS4 的加密算法

5. 解密算法

SMS4 算法的加密解密结构相同，只是密钥的使用次序相反，便于降低实现成本。

6. 密钥扩展

加密密钥为 128 位，需要 32 个 32 位的轮密钥，故需要使用密钥扩展算法。算法结构与加密算法类似，具体细节从略。

经测试，SMS4 可以抵抗差分分析、线性攻击等，且 SMS4 中 S 盒的设计相当安全，在非线性度，自相关性、代数免疫性等方面有相当高的水平。

研究与思考

[1]　比较 WAPI 和 802.11i 的安全性(含认证机制和保密完整性机制)。

[2]　用 C 语言实现 SMS4 算法，并进行性能分析。

[3]　对 SMS4 算法进行安全分析。

进一步阅读建议

2005 年 ACM 安全领域的旗舰会议 CCS 上有一篇论文对 IEEE 802.11i 的安全性进行了证明。

[1]　C. He, M. Sundararajan, A. Datta, A. Derek and J. C. Mitchell. A Modular Correctness Proof of IEEE 802.11i and TLS [C], In Proc. of the 12th ACM conference on Computer and Communications Security (CCS05), 2005.

本章参考文献

[1] IEEE 802.11-1999：Wireless LAN Medium Access Control（MAC）and Physical Layer（PHY）Specifications[S]. 1999.

[2] S. R. Fluhrer，I. Mantin and A. Shamir，Weaknesses in the Key Scheduling Algorithm of RC4 [C]. In Proc. of Selected Areas in Cryptography 2001（SAC01），1-24.2001.

[3] A. Klein，Attacks on the RC4 stream cipher [J]，Designs，Codes and Cryptography，48:269-286，2008.

[4] N. Borisov，I. Goldberg，D. Wagner. Intercepting Mobile Communications：The Insecurity of 802.11 [C]. In Proc. of Seventh Annual International Conference on Mobile Computing and Networking（MOBICOM01），ACM，July 16-21，2001.

[5] 802.1X-2004 - Port Based Network Access Control[S]，http://www.ieee802.org/1/pages/802.1x-2004.html.

[6] Extensible Authentication Protocol（EAP）[S]，RFC 3748，June 2004，http://tools.ietf.org/html/rfc3748.

[7] Remote Authentication Dial In User Service（RADIUS）[S]，RFC 2865，June 2000，http://tools.ietf.org/html/rfc2865.

[8] IEEE 802.11i-2004：Amendment 6：Medium Access Control（MAC）Security Enhancements[S]，IEEE Standards，http://standards.ieee.org/getieee802/download/802.11i-2004.pdf.

[9] 国家商用密码管理办公室，无线局域网产品使用的 SMS4 密码算法[S]，http://www.oscca.gov.cn/UpFile/200621016423197990.pdf.

[10] 国家密码管理局公告(第 7 号)，http://www.oscca.gov.cn/News/200810/News_1104.htm.

[11] 秦兴桥，WAPI 鉴别机制研究与实现[D]，国防科学技术大学硕士论文，2007.

[12] 曹天杰，张永平，汪楚娇. 安全协议[M]. 北京：北京邮电大学出版社，2009.

[13] 马建峰，朱建明等. 无线局域网安全——方法与技术[M]. 北京：机械工业出版社，2005.

[14] 张焕国，刘玉珍. 密码学引论[M]. 武汉：武汉大学出版社，2004.

第6章　远距离无线接入——无线移动通信安全

传感器结点可能通过无线移动通信网络(如GPRS或者TD-SCDMA)直接将收集到的数据传递到中央控制点(例如M2M应用),或者发送至网关后再通过远距离无线移动通信发送到中央控制点(在最终到达中央控制点前可能还需要经过IP核心网)。智能手机结合RFID功能可以实现移动支付(手机钱包)等功能,M2M(如M2M的关键应用远程抄表等)也是由移动通信运营商主推的物联网业务,智能电网也可能利用M2M技术将电力消费(以及电力生成)数据发送到中央控制点。这些都离不开移动通信网络的安全,本章重点介绍2G和3G通信网络中的典型安全问题,即接入认证(鉴权)和数据(保密、完整性)保护机制。

6.1　无线移动通信安全简介

6.1.1　移动通信系统的体系结构

1. 2G/2.5G移动通信系统

这里略去了1G移动通信系统的介绍。2G系统主要采用数字的时分多址(TDMA)和码分多址(CDMA)技术,提供数字化的话音业务及低速数据业务。它克服了模拟移动通信系统的弱点,话音质量和保密性得到很大的提高,并可进行省内、省际自动漫游。具有代表性的2G通信系统有美国的CDMA95系统(基于CDMA技术)和欧洲的GSM系统(基于TDMA技术)。

针对2G系统在数据业务上的弱点,2.5G系统在2G网络中添加分组交换控制功能,可为用户提供一定速率的数据业务(如GPRS系统最大传输速率为115Kbps,CDMA2000 1X系统最大为150Kbps),从而成为介于2G和3G系统的过渡类型。代表性的2.5G系统有基于GSM的GPRS系统和基于CDMA95的CDMA2000 1X系统。

以GSM为例,GSM系统的组成如图6.1所示,主要包括移动台(Mobile Station,MS),基站子系统(Base Station Subsystem),网络子系统(Network Substation,NSS)等几个部分。其中BSS包括基站控制器(Base Station Controller,BSC)和基站收发台(Base Transceiver Station,BTS),网络子系统主要包括移动业务交换中心(Mobile Switch Center,MSC),归属位置寄存器(Home Location Register,HLR),访问位置寄存器(Visitor Location Register),鉴权中心(Authentication Center,AUC),设备识别寄存器(Equipment Identity Register,EIR)等。Um为MS和BTS之间的无线接口。下面分别给予简介。

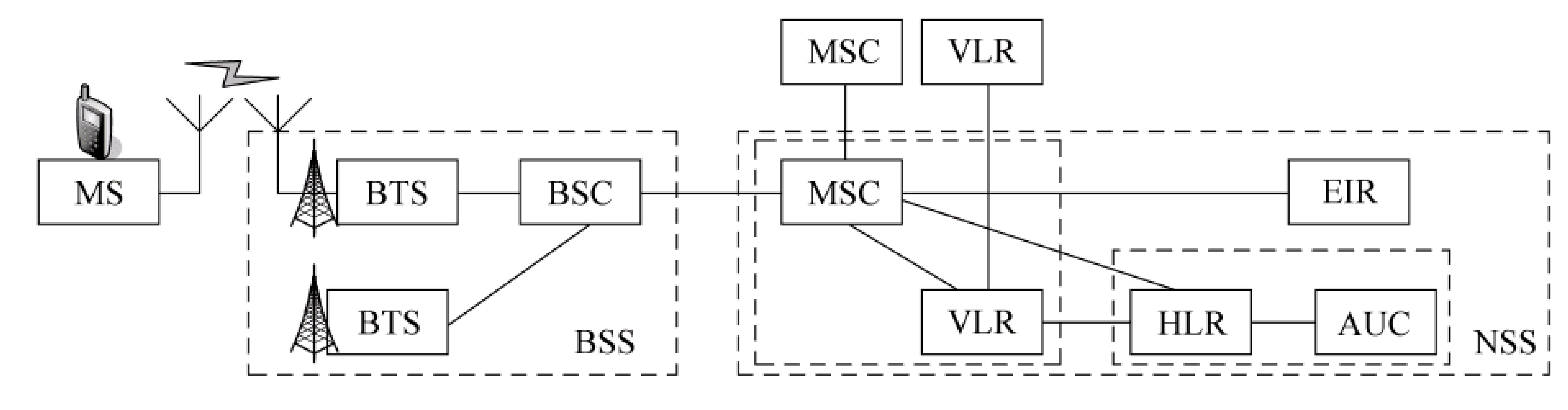


图 6.1　GSM 的系统结构

（1）移动台（MS）。移动台由两部分组成：移动终端和客户识别卡（SIM 卡）。移动终端完成话音编码（信源编码）、信道编码、信息加密、信息的调制和解调、信息发射和接收等功能。SIM 卡存有认证客户身份所需的所有信息，并能执行一些与安全有关的运算，以防止非法客户进入网络，只有插入 SIM 卡后移动台才能接入网内。

（2）基站子系统（BSS）。基站子系统在 GSM 网络的固定部分和移动台之间提供中继。一方面 BSS 通过无线接口直接与移动台实现通信连接，另一方面 BSS 又连接到移动交换子系统（MSS）的移动交换中心（MSC）。BSS 可分为两部分：通过无线接口与移动台相连的基站收发台（BTS）以及与移动交换中心相连的基站控制器（BSC），BTS 负责无线传输、BSC 负责控制与管理。

（3）移动交换中心（MSC）。MSC 是 PLMN（Public Land Mobile Network，公用陆地移动通信网）的核心。MSC 对位于它所覆盖区域中的移动台进行控制和完成话路接续的功能，也是 PLMN 和其他网络之间的接口。它完成通话接续，计费，BSS 和 MSC 之间的切换和辅助性的无线资源管理、移动性管理等功能。另外，为了建立到移动台的呼叫路由，每个 MSC 还完成 GMSC（Gateway Mobile Switching Centre，网关移动交换中心）的功能，即可以查询移动台的位置信息。MSC 从 3 种数据库：访问位置寄存器（VLR）、归属位置寄存器（HLR）和鉴权中心（AUC）中取得处理用户呼叫请求所需的全部数据。反之，MSC 可根据其最新数据更新数据库。

（4）访问位置寄存器（VLR）。VLR 通常与 MSC 在一起，其中存储了 MSC 所管辖区域中的移动台（称访问客户）的相关用户数据，包括：用户号码、移动台的位置区信息、用户状态和用户可获得的服务等参数。VLR 是一个动态用户数据库，它从移动用户的归属位置寄存器（HLR）处获取并存储必要的数据。一旦移动用户离开 VLR 的控制区域，则重新在另一个 VLR 登记，原 VLR 将删除该移动用户的数据记录。

（5）归属位置寄存器（HLR）。HLR 存储管理部门用于移动用户管理的数据。每个移动用户都应在其归属位置寄存器（HLR）注册登记，它主要存储两类信息：有关移动用户的参数，包括移动用户识别号码、访问能力、用户类别和补充业务等数据；有关移动用户目前所处位置的信息，以便建立到移动台的呼叫路由，例如 MSC、VRL 地址等。

（6）鉴权中心（AUC）。AUC 属于 HLR 的一个功能单元，专门用于 GSM 系统的安全性管理。鉴权中心产生鉴权三元组（下一节将介绍）用来鉴权用户身份的合法性以及对无线接口上的话音、数据、信令信号进行加密，防止非授权权用户接入和保证移动用户通信的安全。

（7）设备识别寄存器(EIR)。EIR 存储有关移动台设备参数。完成对移动设备的识别、监视、闭锁等功能，以防止非法移动台的使用。EIR 也是一个数据库，保存着关于移动设备的国际移动设备识别码(IMEI)的三份名单：白名单、黑名单和灰名单。在这三种名单的三种表格中分别列出了准许使用的、出现故障需监视的、丢失不准使用的移动设备的 IMEI 识别码，通过对这三种表格的核查，使得运营部门对于不管是丢失还是由于技术故障或误操作而危及网络正常运营的设备，都能采取及时的防范措施，以确保网络内所使用的设备的唯一性和安全性。

2. 3G 移动通信系统

3G 有更宽的带宽(其传输速度为 384Kbps～2Mbps，带宽可达 5MHz 以上)和系统容量，可实现高速数据传输和多媒体服务。3G 系统的空中接口包含有 WCDMA、CDMA2000 和 TD-SCDMA 3 个标准。其中 WCDMA 是欧洲倡导的宽带 CDMA 技术，该标准提出了 GSM-GPRS-EDGE-WCDMA 的演进方案。而 CDMA2000 标准是美国主推的宽带 CDMA 技术，提出了 CDMA95-CDMA2000 1X-CDMA2000 的演进策略。我国提出的 TD-SCDMA 标准非常适用于 GSM，可以不经过 2.5G 时代，直接向 3G 过渡。和上一章介绍的 WAPI 一样，TD-SCDMA 是我国提出的具有自主知识产权的国际标准(3G 总体架构)。

鉴于 TD-SCDMA 是我国提出的国际标准，这里额外介绍一下。TD-SCDMA(Time Division-Synchronous Code Division Multiple Access)即时分同步码分多址，将 SDMA、同步 CDMA 和软件无线电等当今国际领先技术融合在一起，可以对频率和不同业务灵活搭配，高效率利用频谱等有效资源，加上有 TDMA 和 FDMA 的支持，使得抗干扰能力强，系统容量大。目前 TC 又可分为 TD-SCDMA 和 TD-HSDPA。TD-SCDMA 主要负责提供语音和视频电话等最高下行速率为 384Kbps 的数据业务，而 TD-HSDPA 是一种数据业务增强技术，可以提供 2.8Mbps 的下行速率。

图 6.2 给出了 3G 的网络体系结构，由三个部分组成：移动终端，无线接入网(Radio Access Network, RAN)，核心网(Core Network，CN)。

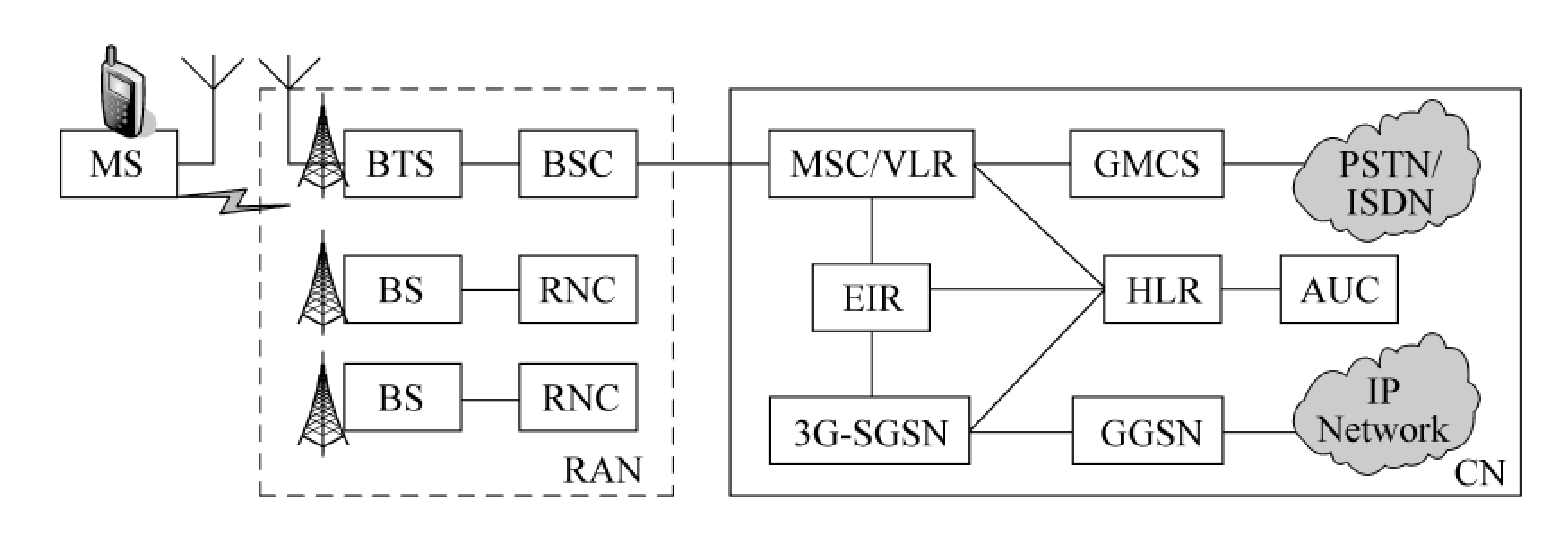


图 6.2　3G 系统结构

（1）移动台。包括两个部分：移动设备(Mobile Equipment，ME)和全球用户识别模块(Universal Subscriber Identity Module，USIM)。ME 实现无线通信功能，USIM 和 SIM 类似，保存了与运营商有关的用户信息。

（2）无线接入网(RAN)。3G 中有两种无线接入网：UTRAN(UMTS Terrestrial

Radio Access Network，UMTS 地面无线接入网络）；GERAN（GSM/EDGE RAN，GSM/EDGE 无线接入网络）。URAN 包括两种网络单元：BS 是 RAN 在网络一侧的终点，BS 被连接到 UTRAN 的控制单元（如无线网络控制器 RNC）上。RNC 通过 Iu 接口与 CN 相连。

（3）核心网（CN）。主要有两个域：分组交换（Packet Switch）域和电路交换（Circuit Switch）域。前者是从 GPRS 域进化而来，其中重要的网络单元是 GPRS 服务支持点 SGSN 和 GPRS 网关支持节点 GGSN，CS 是从传统的 GSM 网络进化而来，其中重要的网络单元是 MSC。SGSN 和 GGSN 是 GPRS 中提出的新网络单元，SGSN 的主要作用是记录移动终端的当前位置信息，且在移动终端和 GGSN 之间完成移动分组数据的发送和接受。GGSN 通过基于 IP 的 GPRS 骨干网连接到 SGSN，是连接 GSM 网络和外部交换网如 Internet 的网关。CN 还可划分为两个部分：本地网络和服务网，本地网络包含有所有有关用户的静态信息，服务网处理用户设备到接入网之间的通信。

6.1.2　移动通信网络的一般安全威胁

按照攻击者攻击的物理位置，对移动通信系统的安全威胁可以分为对无线链路的威胁、对服务网络（有线网络）的威胁和对移动终端（手机）的威胁。由于无线网络的开放性，对无线链路的威胁是需要首先考虑的问题，主要如下：

（1）窃听。由于链路的开放性，在无线链路或服务网内攻击者可以窃听用户数据、信令数据及控制数据，试图解密，或者进行流量分析，即主动或被动流量分析以获取信息的时间、速率、长度、来源及目的地。

（2）伪装。伪基站截取用户数据、信令数据，伪终端欺骗网络获取服务。

（3）用户获取对非授权服务的访问。

（4）破坏数据的完整性。即修改、插入、重放、删除用户数据或信令数据以破坏数据的完整性。

此外，还存在否认攻击，即用户否认业务费用、业务数据来源及发送或接收到的其他用户的数据，网络单元否认提供的网络服务。拒绝服务攻击，在物理上或协议上干扰用户数据和信令数据在无线链路上的正确传输；或使耗尽网络资源，使其他合法用户无法访问。

本章主要从安全协议的角度介绍应对上面所列四种最主要的威胁。应对窃听威胁主要依靠加密的方法；应对伪装主要依靠用户认证的方法；应对破坏数据的完整性的威胁主要依靠采用消息鉴别码的方法。这些防御方法的一个关键点就是需要建立相应的密钥，例如认证密钥、加密密钥、消息完整性密钥。另外，加密密钥通常都是需要经常更换的，以应对敌手对密钥的破解。这种经常更换的用于某个会话的密钥称为会话密钥。虽然在上一章介绍的 WLAN 中接入认证协议与本章将要介绍的接入认证协议了形成类比，但两者的特点有很大不同，如网络的架构显然不同，且移动通信网络中的安全机制更加关注对大规模移动终端的大范围移动性的支持（因为有全球漫游业务、于是在移动通信网络中有宿主网络和被访问网络）。

6.2 2G(GSM)安全机制

6.2.1 GSM的安全需求

蜂窝网(Cellular Network)为移动用户提供无线语音接入,是基于基础设施的网络。基础设施由无线基站和有线骨干网络组成,有线骨干网络将基站连接起来,每一个基站仅在一个有限的物理区域(Cell)内服务。所有基站连接起来便可以覆盖一块很大的区域。手机是典型的蜂窝网终端设备,通过无线信道连接到某个基站。通过基站及其骨干网基础设施,手机可以发起和接收来自其他手机的电话。系统中唯一无线的部分是连接手机和基站的部分,其余的(如骨干网部分)都是有线网络。

GSM(Global System for Mobile Communications,全球移动通信系统)是第二代数字蜂窝移动通信系统的典型例子,其最主要的安全需求是用户的认证接入(因为涉及通信服务计费的问题)。除了用户认证之外,GSM还需要对无线信道内在的威胁(如窃听)采取措施。这样就需要在空中接口上传送的语音通信和传送信息提供保密性。此外,还需要保护用户的隐私,即隐藏用户的真实身份(标识)。

在实际中,GSM安全架构中的一个基本假设就是:用户与宿主网络(Home Network)之间具有长期共享的秘密密钥(存放在SIM卡中),这个秘密密钥是用户认证的基本依据。

6.2.2 GSM用户认证与密钥协商协议

1. 用户认证

在GSM中,秘密密钥和其他与用户身份相关的信息存储在一个安全单元中,称为SIM(Subscriber Identity Module,用户身份识别模块)。SIM以智能卡的形式实现,可以插入手机或者从手机中移除。虽然密钥也可以存储在手机的非易失性存储设备中,用一个口令进行加密,但是将密钥存储在可移除性模块中是一个更好的选择,因为这样允许用户身份在不同的设备之间具有移植性,用户只需要取出SIM卡,便可以更换手机而保持同样的身份。

GSM中的用户认证基于挑战-应答方式,即认证方(网络运营商)提出问题,被认证方(移动终端)进行回答。移动终端收到一个不可预知的随机数作为一个挑战,为了完成认证,必须计算出一个正确的应答。正确的应答必须通过秘密密钥计算得来,不知道秘密密钥便不能计算出正确的应答。因此,如果网络运营商接收到一个正确的应答,就可认为这是一个合法用户。询问必须保证随机性,否则应答就可以预测或者重放。因为挑战值是不可预测的,所以网络运营商在发送挑战后,一定知道应答是刚刚计算的,而不是重放的以前的应答。用于认证的计算由用户的手机和SIM卡完成,无需手机用户的人工参与。

假设用户漫游到一个本地服务区域外的网络,通常称为被访问网络(Visited Network)。GSM用户认证协议的步骤解释如下:

(1) 手机从 SIM 中读取 IMSI(International Mobile Subscriber Identity,国际移动用户识别码),并且将其发送给被访问网络。

(2) 通过 IMSI,被访问网络确定此用户的宿主网络,然后凭借骨干网,被访问网络将 IMSI 转发给用户所在宿主网络。

(3) 宿主网络查询对应于 IMSI 的用户秘密密钥,然后生成一个三元组(RAND,SRES,CK),其中 RAND 是一个伪随机数(Pseudorandom Number),SRES 是此询问的正确应答(Signed RESponse),CK 是用于加密的密钥(即加密会话内容的会话密钥,Cipher Key),用于加密空中接口(手机和被访问网络基站之间的无线通信接口)中传递的内容。RAND 由伪随机数生成器(PseudoRandom Number Generator,PRNG)产生。在 GSM 规范中,SERS 和 CK 为 RAND 和 K 分别利用 A3 算法和 A8 算法(两种专属算法)计算而来。将三元组(RAND,SRES,CK)发送到被访问网络。

(4) 被访问网络向手机发送询问 RAND。

(5) 手机将 RAND 转到 SIM,SIM 计算并且输出应答 SRES′和加密密钥 CK′。手机将 SRES′发送到被访问网络,然后将其与 SERS 比较。如果 SRES′=SRES,则用户得到认证。用户认证成功后,手机和被访问网络的基站之间的通信用会话密钥 CK 加密,容易看出,会话密钥 CK′=CK。使用的加密算法为序列密码算法,在 GSM 规范中叫 A5 算法(安全分析见参考文献,A5/2 是不安全的[1],A5/1[2]和 A5/3,又叫 KASUMI[3])。图 6.3 总结了此认证过程。

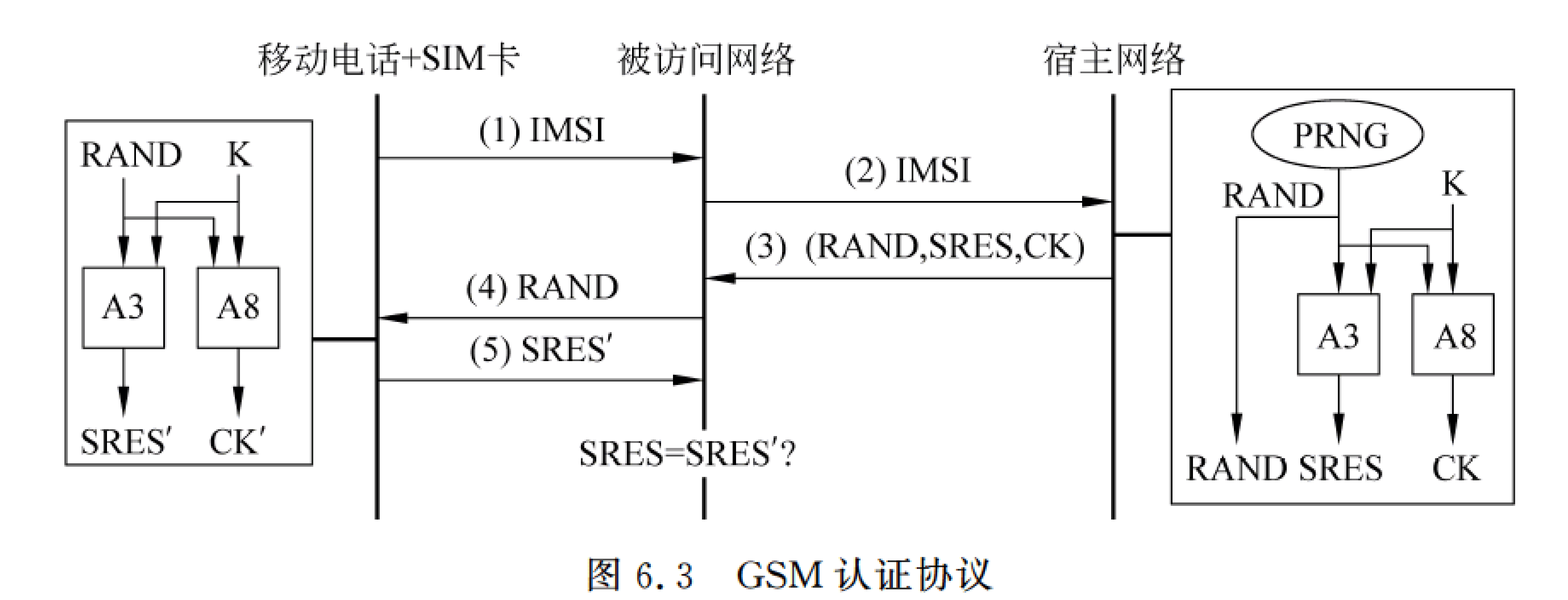


图 6.3　GSM 认证协议

GSM 认证机制虽然与 IEEE 802.11 共享密钥认证类似,但被访问网络和认证网络是分离的,其特点是:被访问网络在不拥有用户长期秘密密钥的情况下,也可以认证用户。即通过宿主网络提供期望的挑战值和相应的应答值给被访问网络来实现。

2. 对用户标识的隐私保护

2G 还提供了对通过认证的用户身份标识 IMSI 的保护,目的是为了保护用户的位置隐私,方法是隐藏空中接口上的 IMSI:用户认证成功后,从被访问网络接收到一个称为 TMSI(Temporary Mobile Subscriber Identity)的临时识别码。TMSI 用新生成的密钥 CK 加密,因而不能被窃听者知道。接下来使用 TMSI,而不是 IMSI。被访问网络保留 TMSI 和 IMSI 间的映射。

当用户进入另外一个被访问网络(不妨称为 B)时,B 联系先前的被访问网络(不妨称为 A),将收到的 TMSI 发送给 A。A 查询与 TMSI 关联的数据并且将用户的 IMSI 和保

留的三元组发送给 B,从而 B 可以为用户服务。如果 TMSI 不再适用(例如手机在漫游到 B 前曾经长时间关机),B 可请求手机发送 IMSI,以重新引导 TMSI 机制。

总之,GSM 安全协议提供了以下的安全服务:

(1) 用户认证基于挑战-应答协议以及用户和宿主网络共享的长期秘密密钥。认证数据从宿主网络发送到被访问网络,长期秘密密钥没有泄漏给被访问网络。

(2) 空中接口(无线链路)上通信的保密性由会话密钥加密来保证,此会话密钥建立在用户认证基础上,在手机和被访问网络间共享,并且由宿主网络协助完成。

(3) 使用临时识别码保护无线接口中的用户真实身份不被窃听识别,即通信中大多数时间不使用真实的身份识别码,窃听者很难追踪用户,从而保护了用户的隐私。

但是 GSM 没有考虑完整性保护的问题,这一点在以语音通信为主的 2G 通信中不是十分重要,因为丢失或者改动的语音通常可以被通话双方人为地识别。但是,由于 3G 中数据业务的增多,必须考虑数据的完整性,因为一个位的改变可能使得数据的含义发生很大改变。完整性问题在接下来介绍的 3G 的安全机制中给予了考虑。

6.3 3G 安全机制

6.3.1 3G 安全体系结构

1. 3G 体系结构

第三代移动通信技术(3G)是指支持高速数据传输的移动通信技术。3G 服务能够同时传送声音及数据信息(电子邮件、即时通信等)。3G 的代表特征是提供高速数据业务,速率一般在几百 Kbps 以上。3G 规范是由国际电信联盟所制定的 IMT-2000 (International Mobile Telecommunications-2000)规范发展而来。

3G 移动通信的主流技术包括 WCDMA、CDMA 2000、TD-SCDMA。WCDMA、TD-SCDMA 的安全规范由欧洲为主体的 3GPP(3G Partnership Project)[4] 制定,其中 TD-SCDMA 由中国提出。CDMA2000 的安全规范由以北美为首的 3GPP2 制定。从某种意义上来讲,WiMAX 也能够提供广域网接入服务。在 2007 年 ITU 将 WiMAX 正式批准为继 WCDMA、CDMA2000 和 TD-SCDMA 之后的第 4 个 3G 标准。鉴于目前 WiMAX (IEEE 802.16)很难在短期内大量应用,本节主要讨论传统从无线移动通信网络演进的 3G 安全。

2. 3G 安全体系结构

3G 系统是在 2G 系统基础上发展起来的,它继承了 2G 系统的安全优点,摒弃了 2G 系统存在的安全缺陷,同时针对 3G 系统的新特性,定义了更加完善的安全特征与安全服务。ETSI TS 133 给出的 3G 安全模型,如图 6.4 所示。3GPP 将 3G 网络划分成了 3 层:应用层、归属层/服务层、传输层。在此基础上将所有安全问题归纳为 5 个范畴:网络接入安全、网络域安全、用户域安全、应用域安全、安全可视性与可配置性。

(1) 网络接入安全。提供安全接入服务网的认证接入机制并抵御对无线链路的窃听篡改等攻击。空中接口的安全性最为重要,因为无线链路最易遭受各种攻击。这一部分

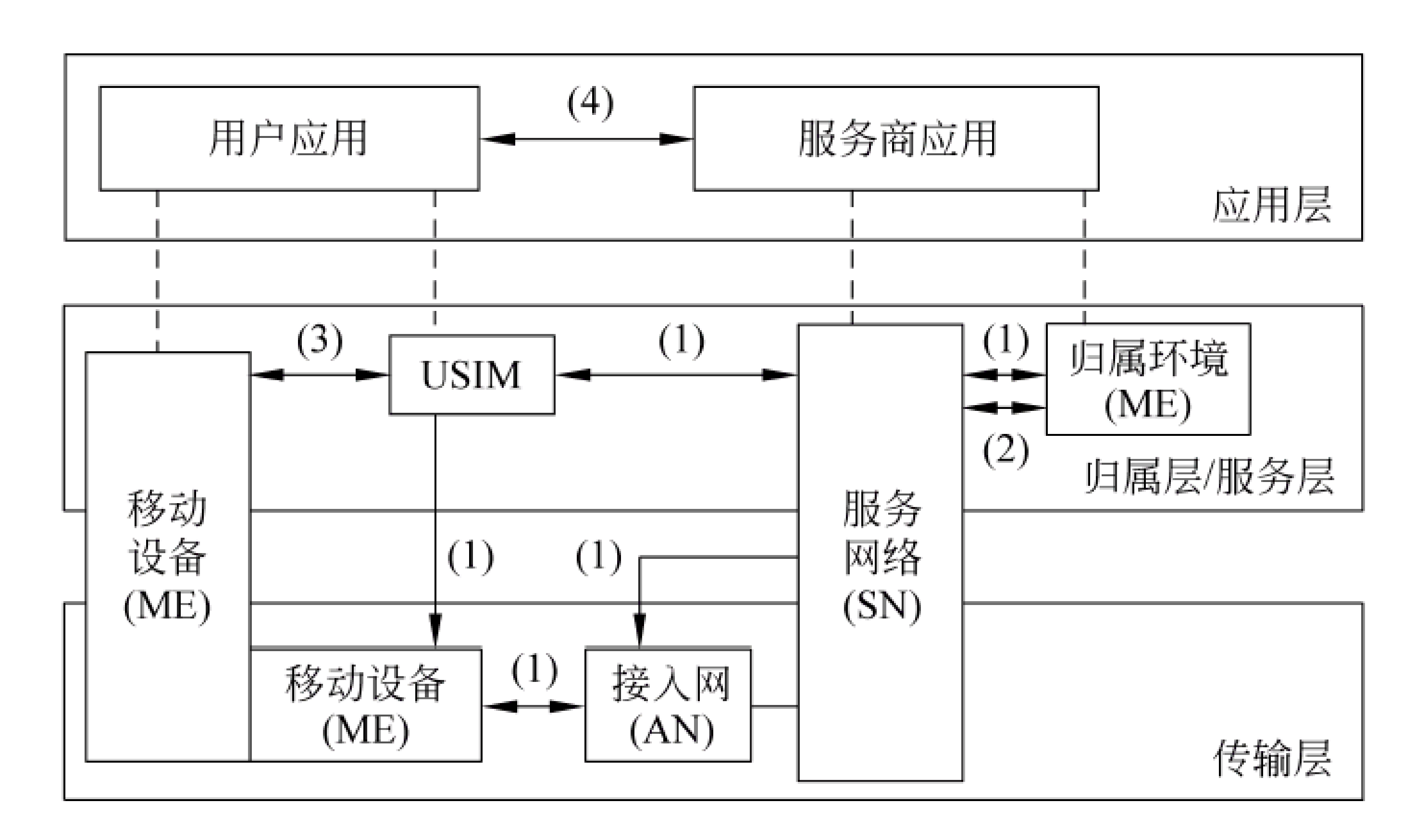


图 6.4　3G 安全总体架构

的功能包括用户身份保密、认证和密钥分配、数据加密和完整性等。其中认证是基于共享对称密钥信息的双向认证，密钥分配和认证一起完成(AKA)。具体包括如下内容。

① 认证：包括对用户的认证和对接入网络的认证。

② 加密：包括加密算法协商、加密密钥协商、用户数据的加密和信令数据的加密。

③ 数据完整性：包括完整性算法协商、完整性密钥协商、数据完整性和数据源认证。

④ 用户标识的保密性：包括用户标识的保密、用户位置的保密及用户位置的不可追踪性，主要是保护用户的个人隐私。

(2) 网络域安全。保证网内信令的安全传送并抵御对有线网络及核心网部分的攻击。网络域安全分为 3 个层次。

① 密钥建立：密钥管理中心产生并存储非对称密钥对，保存其他网络的公钥，产生、存储并分配用于加密信息的对称会话密钥，接收并分配来自其他网络的对称会话密钥。

② 密钥分配：为网络中的结点分配会话密钥。

③ 安全通信：使用对称密钥实现数据加密和数据源认证。

(3) 用户域安全。用户服务识别模块(User Service Identity Module，USIM)是一个运行在可更换的智能卡上的应用程序。用户域安全机制用于保护用户与用户服务识别模块之间，以及用户服务识别模块与终端之间的连接，包括以下两个部分。

① 用户到用户服务身份模块(USIM)的认证：用户接入到 USIM 之前必须经过 USIM 的认证，确保接入到 USIM 的用户为已授权用户。

② USIM 到终端的连接：确保只有授权的 USIM 才能接入到终端或其他用户环境。

(4) 应用域安全。用户域与服务提供商的应用程序间能安全地交换信息。USIM 应用程序为操作员或第三方运营商提供了创建驻留应用程序的能力，需要确保通过网络向 USIM 应用程序传输信息的安全性，其安全级别可由网络操作员或应用程序提供商根据需要选择。

(5) 安全特性的可视性及可配置能力。安全特性的可视性指用户能获知安全特性是否正在使用，服务提供商提供的服务是否需要以安全服务为基础。确保安全功能对用户来说是可见的，这样用户就可以知道自己当前的通信是否已被安全保护、受保护的程度是

多少。例如，接入网络的加密提示，通知用户是否保护传输的数据，特别是建立非加密的呼叫连接时进行提示；安全级别提示通知用户被访问网络提供什么样的安全级别，特别是当用户漫游到低安全级别的网络（如从 3G 到 2G）时进行提示。可配置性指允许用户对于当前运行的安全功能进行选择配置，包括是否允许用户到 USIM 的认证；是否接收未加密的呼叫；是否建立非加密的呼叫；是否接受某种加密算法。

从以上分析可知，3G 网络安全的特殊性在于添加了对用户域和网络域安全的考虑；安全特性的可视性和可配置能力体现出对用户参与性的考虑；应用域安全表现了对 USIM 应用程序复杂性的考虑。由于网络接入安全在 3G 安全中具有重要的地位，下面主要介绍 3G 接入安全的认证与密钥协商协议。

6.3.2 3G(UMTS)认证与密钥协商协议

通用移动通信系统（Universal Mobile Telecommunications System，UMTS）是当前使用最广泛的一种 3G 移动通信技术，它的无线接口使用 WCDMA 和 TD-SCDMA。UMTS 从 GSM 到 GPRS（2.5G）演进（GPRS 支持更好的数据速率，理论上最大可以到 140.8Kbps，实际上能实现接近 56Kbps，已经在很多 GSM 网络部署，它也是目前很多 M2M 应用所采用的技术）而来，故两者的系统的架构十分相似。UMTS 提供的接入安全是 GSM 相关安全特性的超集，它相对于 GSM 的新安全特性是用于解决 GSM 中潜在的安全缺陷。

1. UMTS 的认证向量

首先回顾 GSM 安全中的主要问题，包括：

（1）单向认证，即只认证接入用户，没有认证被访问网络。

（2）GSM 认证三元组可无限期使用。认证协议中用户无法验证接收到的挑战是否新鲜。

（3）空中接口上的通信和传输没有完整性保护服务。

（4）加密密钥长度太短。用户的长期密钥可能泄漏，SIM 卡可能被克隆。

UMTS 的安全架构解决了上述安全问题，但是为了尽可能地兼顾原有的设备投资，设计在 GSM 安全架构基础之上进行了扩展，GSM 中的三元组被替换为认证五元向量：（RAND，XRES，CK，IK，AUTN）。和 GSM 中一样，RAND 是一个不可预知的伪随机数，由 PRNG 产生，并且在认证协议中作为挑战，XRES 为 RAND 的期望应答（eXpected RESponse），CK 是会话加密密钥（Cipher Key）。XRES 和 CK 都由 RAND 和用户的长期秘密密钥 K 计算得来。此外，IK 为完整性保护密钥（Integrity Key），AUTN 是一个认证令牌（Authentication Token），用于给用户提供对宿主网络的认证，并且保证了 RAND 的新鲜度。AUTN 由 3 个字段组成：

$$AUTN=(SQN \oplus AK,\ AMF,\ MAC)$$

这里 SQN 是一个由用户和宿主网络动态维持的序号（Sequence Number）；AI 称为匿名密钥（Anonymity Key），用于保护 SQN 以防窃听者偷听；AK 由 RAND 和 K 产生；AMF 是认证管理字段（Authentication Management Field），用于在宿主网络和用户之间传递参数；MAC 是一个消息鉴别码，在 RAND、SQN 和 AMF 上利用长期密钥 K 计算

而来。

AUTN 的结构和认证向量在图 6.5 中进行了直观描述。函数 f1、f2、f3、f4 和 f5 为 UMTS 标准中定义的单向函数。PRNG 是伪随机数生成器。

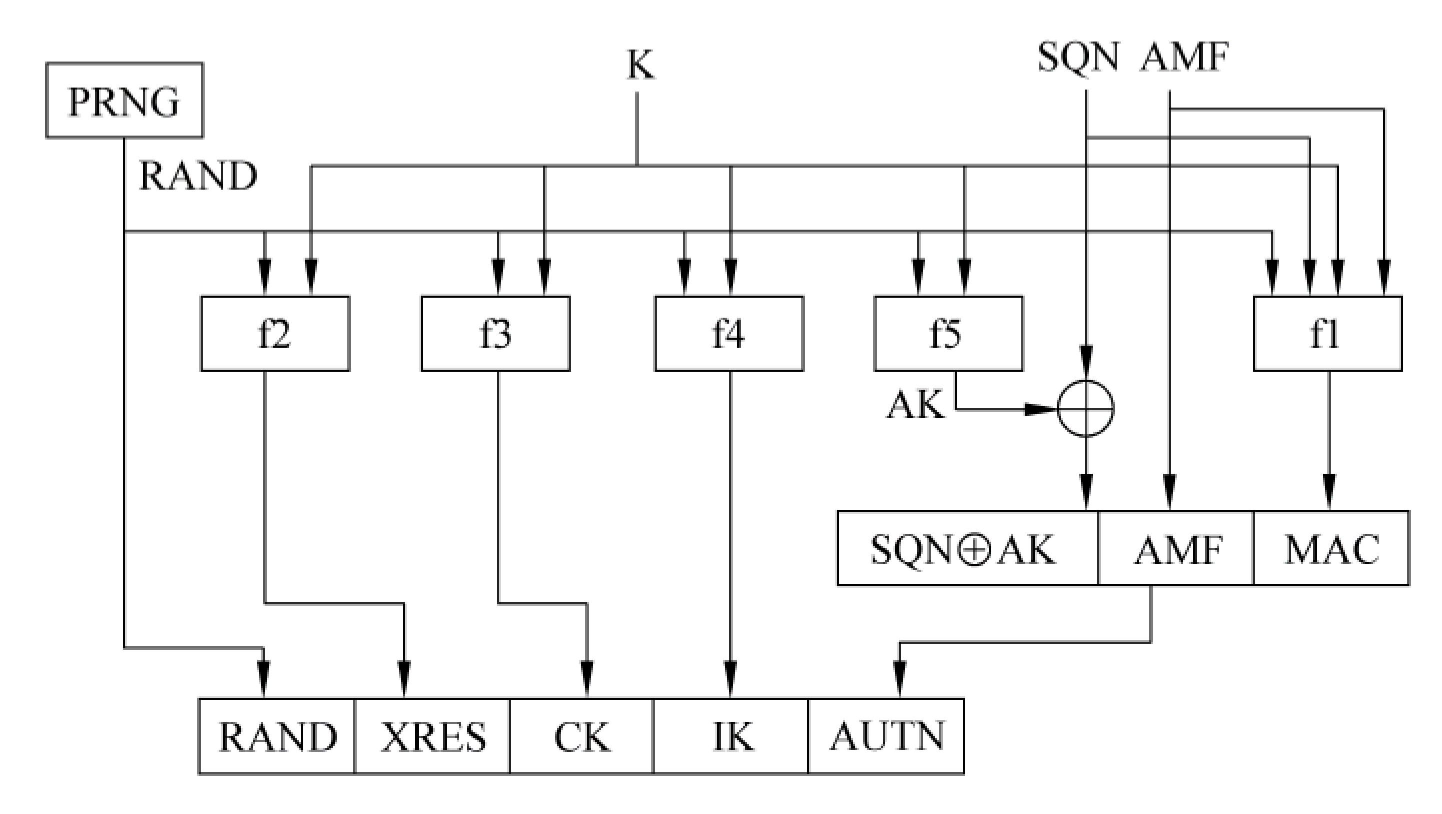


图 6.5　UMTS 中的认证向量及 AUTN 令牌的构成

2. 3G 接入认证与密钥协商协议

3G 网络中的接入安全要确保的内容包括两部分：提供用户和网络之间的身份认证，以保证用户和网络双方的实体可靠性；空中接口安全，主要用于保护无线链路传输的用户和信令信息不被窃听和篡改。前者需要身份认证，后者需要加密和消息完整性保护，而这些离不开密钥，因此需要进行密钥的协商。安全协议设计时，通常将两者一起考虑，以提高协议的效率，减少消息交换的次数，即先进行认证，然后进行密钥协商，这一协议机制称为认证密钥协商机制（Authentication and Key Agreement，AKA）。虽然在前面的章节曾经多次提到过，但这里的区别是对移动性的支持。

3G 认证与密钥协商协议（3G AKA）中参与认证和密钥协商的主体有用户终端（ME/USIM）、被访问网络（Visitor Location Register/Servicing GPRS Support Node，VLR/SGSN）和归属网络（Home Environment/Home Location Register，HE/HLR）。在 3G AKA 协议中，通过用户认证应答（RES）实现 VLR 对 ME 的认证，通过消息鉴别码（MAC）实现 ME 对 HLR 的认证，以及实现了 ME 与 VLR 之间的密钥分配，同时每次使用的消息鉴别码 MAC 是由不断递增的序列号（SQN）作为其输入变量之一，保证了认证消息的新鲜性，从而确保了密钥的新鲜性，有效地防止了重放攻击。3G 认证和密钥协商过程如下（如图 6.6 所示）。

（1）移动终端（ME/USIM）向网络发出呼叫接入请求，把身份标识（IMSI）发给 VLR。

（2）VLR 收到该注册请求后，向用户的 HLR 发送该用户的 IMSI，请求对该用户进行认证。

（3）HLR 收到 VLR 的认证请求后，生成序列号 SQN 和随机数 RAND，计算认证向量 AV 发送给 VLR。其中，AV=RAND‖XRES‖CK‖IK‖AUTN。K 为 ME 和 HLR 共同拥有的永久性密钥，写入在 ME 中的 SIM 卡中。AV 各字段的计算方法如下（如图 6.7 所示）。

① XRES=$f2_K$(RAND)，期望的应答（eXpected RESponse）。

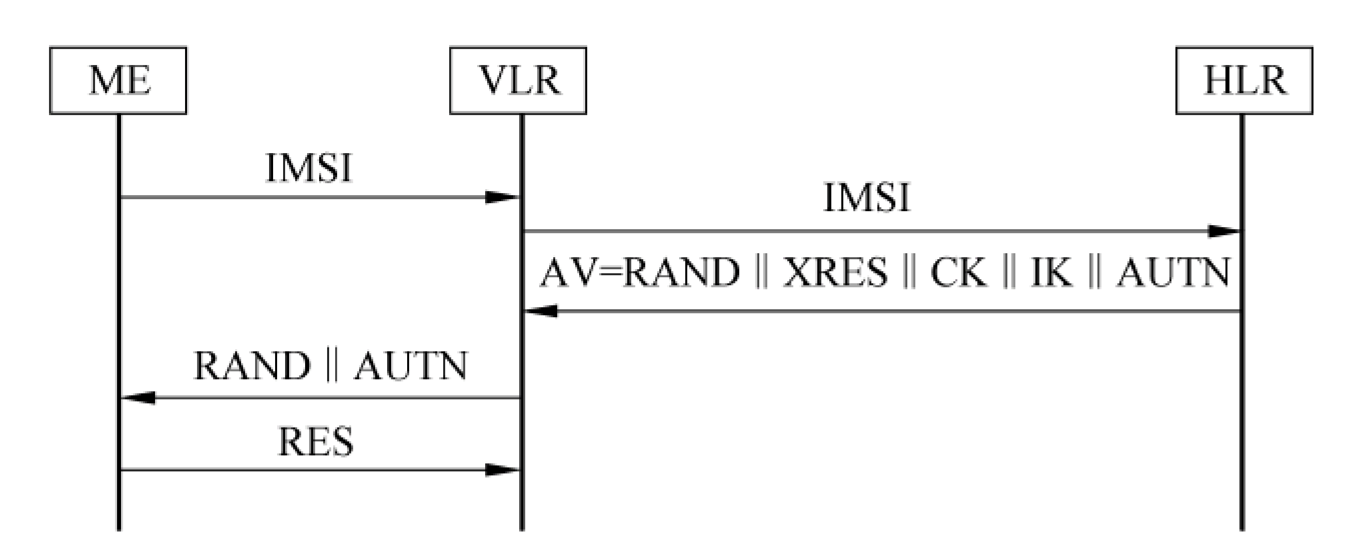


图 6.6　3G 认证和密钥协商(AKA)过程

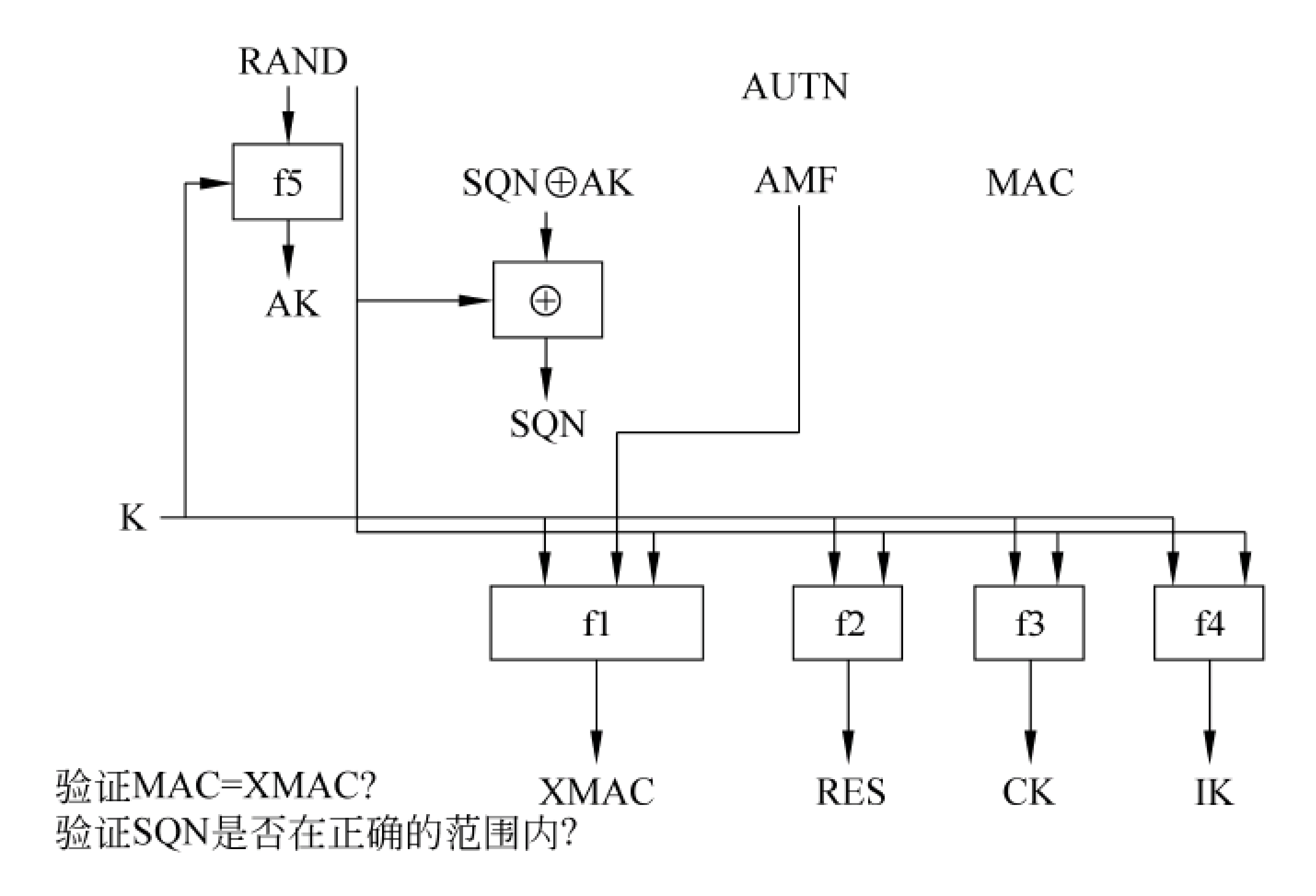


图 6.7　生成认证向量 AV 的过程

② CK＝$f3_K$(RAND),加密密钥;IK＝$f4_K$(RAND),完整性密钥。

③ AUTN＝SQN⊕AK ‖ AMF ‖ MAC,认证令牌。

④ SQN:序列号。

⑤ AK＝$f5_K$(RAND),匿名密钥,用于隐蔽序列号。

⑥ AMF:鉴别管理字段(Authentication Management Field)。

⑦ MAC＝$f1_K$(SQN ‖ RAND ‖ AMF),消息鉴别码。

这里 f1 算法用于产生消息鉴别码,f2 算法用于接入认证中计算期望的应答值。f3、f4、f5 是密钥生成函数,f3 算法用于产生加密密钥,f4 算法用于产生完整性密钥,f5 算法用于产生匿名密钥,5 个函数的具体内容由 3GPP 相关规范给出。这里 f_K 表示函数 f 使用 K 作为密钥。表 6.1 总结了 AKA 中使用到的 5 个函数。

表 6.1　3G 接入安全中使用的部分主要函数

函　数　名	函 数 用 途	函 数 输 入	函 数 输 出
f0	随机数生成函数	无	RAND
f1	消息鉴别函数	K, SQN, RAND, AMF	XMAC/MAC
f2	生成期望的应答的鉴权函数	K, RAND	XRES/RES

续表

函　数　名	函 数 用 途	函 数 输 入	函 数 输 出
f3	加密密钥生成函数	K，RAND	CK
f4	完整性密钥生成函数	K，RAND	IK
f5	匿名密钥生成函数	K，RAND	AK

(4) VLR 接收到认证向量后，将 RAND 及 AUTN 发送给 ME，请求用户产生认证数据。

(5) ME 接收到认证请求后，首先计算 XMAC，并与 AUTN 中的 MAC 比较，若不同，则向 VLR 发送拒绝认证消息，并放弃该过程。同时，ME 验证接收到的 SQN 是否在有效的范围内，若不在有效的范围内，ME 则向 VLR 发送“同步失败”消息，并放弃该过程。上述两项验证通过后，ME 计算 RES、CK 和 IK，并将 RES 发送给 VLR。

因为 ME 和 HLR 都预先知道相同的计算方法，因此 XMAC、RES 计算如下：

消息鉴别码：$XMAC=f1_K(SQN \| RAND \| AMF)$

用户认证应答：$RES=f2_K(RAND)$

(6) VLR 接收到来自 ME 的 RES 后，将 RES 与认证向量 AV 中的 XRES 进行比较，若相同则 ME 的认证成功，否则 ME 认证失败。最后，ME 和 VLR 建立的共享加密密钥是 CK，数据完整性密钥是 IK。

6.4　4G 安全机制简介

6.4.1　4G 国际标准 TD-LTE-A

4G 通信将能满足 3G 不能达到的覆盖范围、通信质量、高速传输率和高分辨率多媒体服务，通常也被称为“多媒体移动通信”。4G 的数据传输率可达到 10～20Mbps，最高甚至达到 100Mbps。其中 TD-LTE 下行速率 100Mbps，上行速率 50Mbps。

3GPP LTE(Long Term Evolution，长期演进)包括两种制式：LTE TDD(时分双工)和 LTE FDD(频分双工)。其中 LTE TDD 就是 TD-LTE 技术，它吸纳了很多 TD-SCDMA 的技术元素，拓展了 TD-SCDMA 在智能天线、系统设计等方面的关键技术和我国自主知识产权，具有高效益低时延、高带宽低成本等特点和优势，系统能力与 LTE FDD 相当。TD-LTE-Advanced(TD-LTE-A)技术方案已经与 2010 年 10 月被国际电信联盟 ITU 确定为 4G 的 2 个国际标准之一，在未来大规模使用具有自主知识产权的 TD-LTE-A 标准发展 4G 对我国将具有极其重要的战略意义。

据 2012 年 2 月人民网的消息，TD-LTE 产业链已受到包括中国移动、日本软银、沃达丰、德国电信等一批国内外主流运营商的认可和接纳，在全球已建设 33 个试验网，预计 2012 年将有超过 10 个国家和地区开始 TD-LTE 网络的商用部署。中国移动 TD-LTE 规模试验网部署项目采取“6＋1”方案，将投资 15 亿人民币建网覆盖上海、杭州、南京、广州、深圳、厦门 6 个城市，每个城市将部署约 200 个基站；并在北京建 TD-LTE 演示网。

*6.4.2 LTE 中的流密码算法 ZUC

ZUC 算法(祖冲之密码算法,ZUC 是我国古代数学家祖冲之名字的缩写)是中国通信标准协会 CCSA(China Communications Standards Association)推荐给 3GPP LTE 使用的新算法,目前 ZUC 算法已通过了算法标准组 ETSI SAGE 的内部评估,ETSI SAGE 认为该算法强壮,并推荐在 LTE 标准中使用。ZUC 算法目前处于公开评估阶段。ZUC 算法是第一个成为国际标准的我国自主知识产权的密码算法,是我国商用密码算法首次走出国门,具有重要的历史意义。ZUC 算法的国际标准化,对我国按照国际惯例掌握通信产业的主动权有非常重要的意义。

ZUC 算法由中国科学院数据保护和通信安全研究中心(DACAS)研制。LTE 算法的核心是 ZUC 算法;由 ZUC 定义的 LTE 加密算法,称为 128-EEA3;由 ZUC 定义的 LTE 完整性保护算法,称为 128-EIA3。

ZUC 是一个面向字(Word-Oriented)的流密码算法。输入是 128 位的初始密钥和 128 位的初始向量,输出是一个 32 位字(也称为密钥字,Key-word)。这一密钥字可用于加密。ZUC 的执行有两个阶段:初始阶段和工作阶段。在初始阶段,密钥/IV 将初始化,耗费时钟周期但没有产生输出;在工作阶段,每个时钟周期都会输出一个 32 位的字。

1. ZUC 算法的整体架构

ZUC 具有 3 个逻辑层[13,14],如图 6.8 所示。最上层是一个具有 16 阶段的线性反

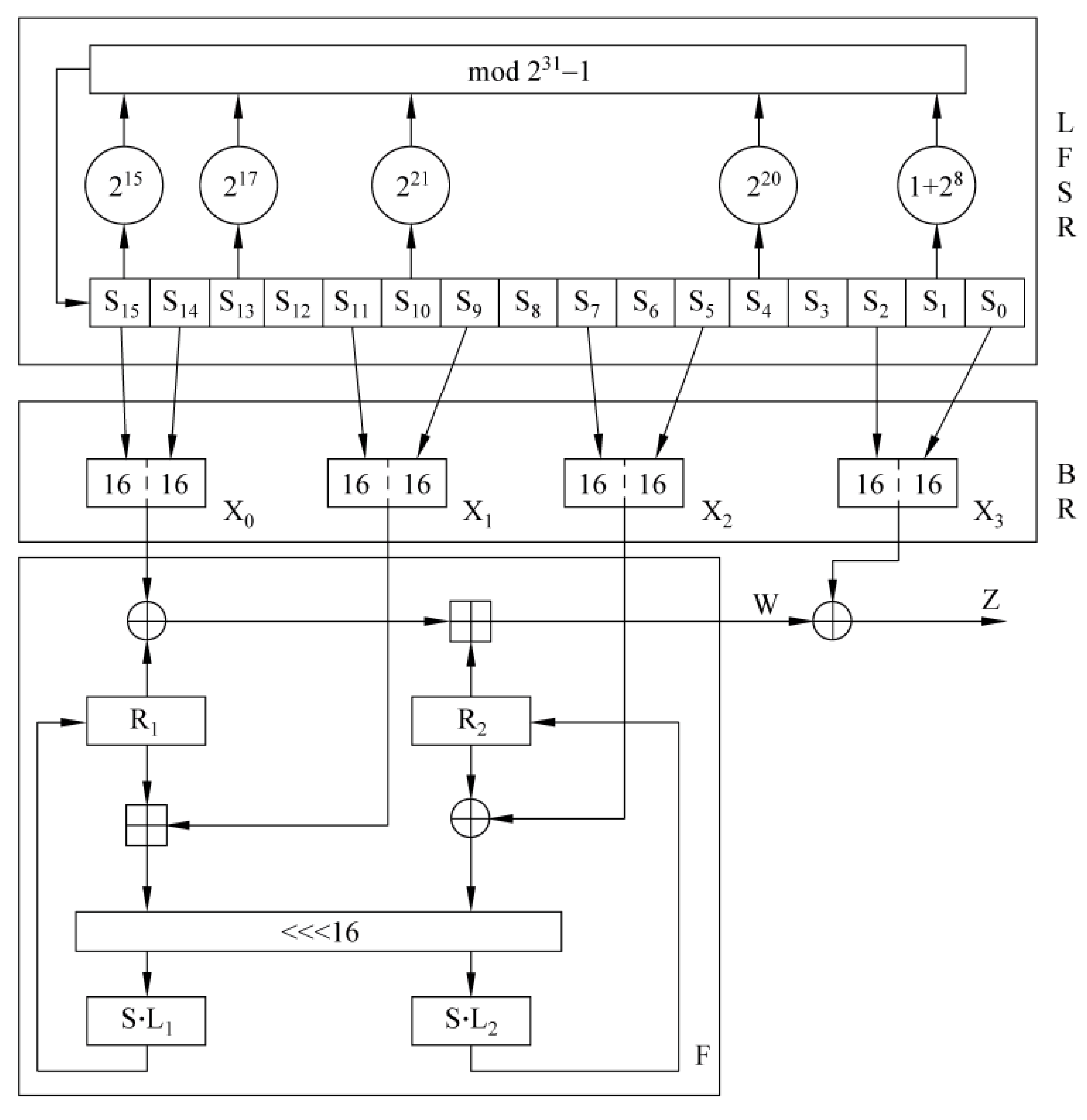


图 6.8 ZUC 算法整体架构

馈移位寄存器(Linear Feedback Shift Register,LFSR),中间层是一个比特混淆器(Bit-Reorganization,BR),最下层是一个非线性函数 F。

2. LFSR

LFSR 有 16 个 31 位的单元($s_0, s_1, \cdots, s_{15}$),每个单元 s_i($0 \leqslant i \leqslant 15$)的输入在集合{1, 2, 3, …, $2^{31}-1$}中。LFSR 具有 2 个操作模式:初始化模式和工作模式。

初始化模式中,LFSR 的输入是 31 位的字 u,该字来自于非线性函数 F 的 32 位的输出 W,将 W 的最右一个位去掉,即 $u = W >> 1$。具体而言,初始化模式为

```
LFSRWithInitialisationMode(u) {
1. v = 2^15 s15 + 2^17 s13 + 2^21 s10 + 2^20 s4 + (1+2^8) s0 mod (2^31 - 1);
2. s16 = (v+u) mod (2^31 - 1);
3. If s16 = 0, then set s16 = 2^31 - 1;
4. (s1, s2, …, s15, s16) → (s0, s1, …, s14, s15).
}
```

工作模式下,LFSR 没有输入,计算方法如下:

```
LFSRWithWorkMode() {
1. s16 = 2^15 s15 + 2^17 s13 + 2^21 s10 + 2^20 s4 + (1+2^8) s0 mod (2^31 - 1);
2. If s16 = 0, then set s16 = 2^31 - 1;
3. (s1, s2, …, s15, s16) → (s0, s1, …, s14, s15).
}
```

上述过程第一步中的 31 位串 s 在 GF($2^{31}-1$)乘以 2^i 可通过 s 的循环左移 i 位来实现,因而只需要模加运算,于是第一步可用如下方法实现:

$$v = (s_{15} <<<_{31} 15) + (s_{13} <<<_{31} 17) + (s_{10} <<<_{31} 21) + (s_4 <<<_{31} 20) + (s_0 <<<_{31} 8) + s_0 \bmod (2^{31}-1)$$

标准中同样对模加给出一些注解。

3. BR 操作

中间层的运算是 BR,它从 LFSR 的单元中提取 128 位,形成 4 个 32 位字,前 3 个字用于最下层的非线性函数 F,最后一个字用于生成密钥流。

令 $s_0, s_2, s_5, s_7, s_9, s_{11}, s_{14}, s_{15}$ 是 LFSR 的 8 个单元,那么 BR 如下生成 4 个 32 位字 X_0, X_1, X_2, X_3:

```
Bitreorganization() {
1. X0 = s15H || s14L;
2. X1 = s11L || s9H;
3. X2 = s7L || s5H;
4. X3 = s2L || s0H.
}
```

由于 s_i 是 31 位整数,于是 s_{iH} 表示 s_i 的 30…15 比特位而不是 31…16 位,这里 $0 \leqslant i$

$\leqslant 15$。

4. 非线性函数 F

非线性函数 F 有 2 个 32 位存储单元 R_1 和 R_2。令 F 的输入是 X_0，X_1 和 X_2，即 BR 操作的输出，F 输出一个 32 位字 W。如下：

F（X_0，X_1，X_2）{
1. $W = (X_0 \oplus R_1) \boxplus R_2$；
2. $W_1 = R_1 \boxplus X_1$；
3. $W_2 = R_2 \oplus X_2$；
4. $R_1 = S(L_1(W_{1L} \| W_{2H}))$；
5. $R_2 = S(L_2(W_{2L} \| W_{1H}))$.
}

S 是一个 32×32 的 S 盒，L_1 和 L_2 线性变换。下面分别加以介绍。

S 是 32 位输入，32 位输出。其实由 4 个并列 8×8 的 S 盒组成，即 $S = (S_0, S_1, S_2, S_3)$，其中 $S_0 = S_2$，$S_1 = S_3$，即实际上定义了 2 个输入输出表 S_0，S_1。两个表的定义见标准文档。例如求 $S_0(x)$，可将 x 写为 $h \| l$，查 S_0 表中 h 行 l 列，$S_0(0 \times 12) = 0xF9$ 以及 $S_1(0 \times 34) = 0xC0$。

线性变换 L_1 和 L_2 输入是 32 位字，输出是 32 位字，定义如下：

$$L_1(X) = X \oplus (X <<<_{32} 2) \oplus (X <<<_{32} 10) \oplus (X <<<_{32} 18) \oplus (X <<<_{32} 24),$$
$$L_2(X) = X \oplus (X <<<_{32} 8) \oplus (X <<<_{32} 14) \oplus (X <<<_{32} 22) \oplus (X <<<_{32} 30)$$

5. 密钥装载

密钥装载将扩展初始密钥和初始向量为 16 个 31 位的整数，作为 LFSR 的初始状态。令 128 位的初始密钥 $k = k_0 \| k_1 \| k_2 \| \cdots \| k_{15}$ 和 128 位初始向量 $iv = iv_0 \| iv_1 \| iv_2 \| \cdots \| iv_{15}$。这里 k_i 和 iv_i 均为字节（$0 \leqslant i \leqslant 15$）。k 和 iv 装载到 LFSR 的 s_0，s_1，…，s_{15} 中：

令 D 为 240 位固定串，由 16 个 15 位的子串组成：$D = d_0 \| d_1 \| \cdots \| d_{15}$，其中

$d_0 = 100010011010111_2$，$d_1 = 010011010111100_2$，$d_2 = 110001001101011_2$，
$d_3 = 001001101011110_2$，$d_4 = 101011110001001_2$，$d_5 = 011010111100010_2$，
$d_6 = 111000100110101_2$，$d_7 = 000100110101111_2$，$d_8 = 100110101111000_2$，
$d_9 = 010111100010011_2$，$d_{10} = 110101111000100_2$，$d_{11} = 001101011110001_2$，
$d_{12} = 101111000100110_2$

然后给 s_i 赋值，令 $s_i = k_i \| d_i \| iv_i$（$0 \leqslant i \leqslant 15$）。

6. ZUC 的执行

ZUC 有两个阶段：初始化阶段和工作阶段。初始化阶段算法调用密钥装载过程，将 128 位初始密钥 k 和 128 位初始向量 iv 装载到 LFSR 中，令 32 位存储单元 R_1 和 R_2 全为 0。于是运行如下操作 32 次：

1. Bitreorganization()；
2. $w = F(X_0, X_1, X_2)$；
3. LFSRWithInitialisationMode($w >> 1$)

初始阶段完成后，算法进入到工作阶段。在工作阶段，算法执行如下操作一次，丢弃掉 F 的输出 W：

1. Bitreorganization()；
2. $F(X_0, X_1, X_2)$；
3. LFSRWithWorkMode()

于是，算法准备生成密钥流，即对下面的操作每执行一次，一个 32 位字 Z 便输出：

1. Bitreorganization()；
2. $Z = F(X_0, X_1, X_2) \oplus X_3$；
3. LFSRWithWorkMode()

该标准文档还给出了 ZUC 算法的 C 语言实现源代码。

研究与思考

[1] 实现 ZUC 算法并进行性能分析。
[2] 对 ZUC 算法进行安全分析。
[3] 思考物联网在网络异构和网络融合环境下的安全问题。

进一步阅读建议

物联网在异构网络和融合网络的安全问题值得注意。ZUC 的密码学分析也处于公开评估阶段。

[1] 李亚晖．异构无线网络安全协议研究[D]．西安电子科技大学博士学位论文，2009.
[2] 马文静．下一代无线网络安全及切换机制研究[D]．北京邮电大学博士学位论文，2010.

本章参考文献

[1] I. Goldberg, D. Wagner, Lucky Green. The (Real-Time) Cryptanalysis of A5/2[C]. Rump session of Crypto'99, 1999.
[2] E. Barkan, E. Biham, N. Keller, Instant Ciphertext-Only Cryptanalysis of GSM Encrypted Communication[C], Crypto 2003：600-16.
[3] O. Dunkelman, Nathan Keller, Adi Shamir, A Practical-Time Attack on the A5/3 Cryptosystem Used in Third Generation GSM Telephony [OL]. http://eprint.iacr.org/2010/013.
[4] 3GPP, http://www.3gpp.org.
[5] 姜奇，马建峰，李光松，马卓．基于 WAPI 的 WLAN 与 3G 网络安全融合[J]．计算机学报，2010(9).
[6] 边瑞昭，刘曼华，王惠芳，马自堂．3G 中安全增强的 AKA 协议设计与分析[J]．计算机应用与软件[J]，2010(1).
[7] 曹晨磊，张茹，钮心忻，周琳娜，张振涛．3G 实体认证协议及技术规范的综述与安全分析[J]．清

华大学学报(自然科学版)，2009/S2.
[8] 戴沁芸. 第三代移动通信系统网络接入安全机制分析[J]. 现代电信科技，2010(4).
[9] 戴沁芸，杨海军，姚钦锋，张春辉. 浅析下一代移动通信网络的安全问题[J]. 信息安全与通信保密，2009 (9) .
[10] 朱红儒，肖国镇. 基于整个网络的 3G 安全体制的设计与分析[J]. 通信学报，2002，23(4).
[11] 张文芳，何大可. 第三代移动通信系统网络接入安全策略[J]. 通信技术，2003(1).
[12] L. Buttyan and J. P. Hubaux，Security and Cooperation In Wireless Networks [M]，A Graduate Textbook，Cambridge University Press，2007.
[13] Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3 Revised versions published July 2011，Document 2：Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3：ZUC Specification[S]，http://gsmworld. com/documents/EEA3_EIA3_ZUC_v1_6. doc.
[14] Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3 Revised versions published July 2011，Document 1：Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3：128-EEA3 & 128-EIA3 Specification[S]，http://gsmworld. com/documents/EEA3_EIA3_specification_v1_6. doc.

第 7 章　接入网安全的扩展讨论

除了第 5 章和第 6 章介绍的近距离无线(高速)接入和远距离无线(高速)接入以外,物联网的接入技术还可以是近距离无线低速接入,如蓝牙和 ZigBee。以及近距离有线接入,如局域网 Ethernet(IEEE 802.3)和现场总线。最后,本章对具有自主知识产权的手机电视协议 CMMB 系统和北斗卫星导航定位系统及其安全进行了简介,以促进对自主创新的关注和二次开发。手机电视以及卫星定位导航系统也都可视为物联网中的关键应用。

7.1　近距离无线低速网络安全

第 5 章和第 6 章分别介绍了近距离无线(高速)接入和远距离无线(高速)接入的典型代表,本节介绍近距离无线低速接入方法的典型代表 Bluetooth 和 ZigBee。此外,红外技术也可以进行无线通信,但限制是只能在两个实体间通信,且有视距(Line of Sight,LOS)的限制,因此不再赘述。

7.1.1　Bluetooth 安全简介

蓝牙(Bluetooth)是爱立信公司在 1994 年开始研究的一种能使手机与其附件(如耳机)之间互相通信的无线模块。它的工作频率为 2.4GHz,有效范围大约在 10 米半径内。蓝牙被列入了 IEEE 802.15.1 规范,规定了包括 PHY、MAC、网络和应用层等集成协议栈。蓝牙技术可解决小型移动设备间的短距离无线互连问题,它的硬件涵盖了局域网中的各类数据及语音设备,如计算机、移动电话等。安全性是整个蓝牙协议中非常重要的一部分,协议在应用层和链路层均提供了安全措施。

1. 蓝牙协议

蓝牙标准包括两大部分:核心(Core)部分定义了蓝牙的技术细节;轮廓(Profile)部分定义了在蓝牙的各种应用中协议栈的组成。蓝牙标准主要定义的是底层协议,同时为保证和其他协议的兼容性,也定义了一些高层协议和相关接口。从 ISO 的 OSI 七层协议标准来看,蓝牙标准主要定义的是物理层、链路层和网络层的结构。

蓝牙的特点包括:

(1) 全球范围使用:蓝牙工作在 2.4GHz 的 ISM(Industry、Science、Medicine)频段,全球大多数国家 ISM 频段范围是 2.4～2.4835GHz,使用该频段无须向各国的无线电资源管理部门申请许可证。

(2) 可以传输语音和数据:蓝牙同时采用了电路交换和分组交换两种交换技术,支

持数据和语音的同时传输。蓝牙中定义了两种链路类型,分别用来传输数据和语音。

(3) 组网灵活：根据蓝牙设备在网络中的角色不同,可以将其分为主设备(Master)和从设备(Slave)。在建立连接时,主动发起连接请求的为主设备,响应方为从设备。几个蓝牙设备可以连接建立成一个微微网(Piconet),在一个微微网中只能有一个主设备,一个主设备最多可以带 7 个从设备。

(4) 体积小,便于集成到其他设备中：蓝牙芯片的封装尺寸已经缩小至不到 16mm。

(5) 功耗低：蓝牙设备在连接状态下,有 4 种工作模式：激活(Active)、呼吸(Sniff)、保持(Hold)和休眠(Park)。Active 模式是正常的工作状态;Sniff 模式下从设备周期性的被激活;Hold 模式下从设备停止监听来自主设备的数据分组;Park 模式下主从设备仍然保持同步,但从设备已经不需要保留其激活成员地址。后 3 种模式为节能模式,定义这 3 种节能模式是为了尽可能降低蓝牙的功耗。

另外,蓝牙的抗干扰能力强,采用跳频方式扩展频谱来避免工作在 ISM 频段的设备干扰,且成本低廉、广泛应用于大量手机和笔记本电脑中。

2. Bluetooth 链路层安全简介

蓝牙的安全模式包括：

(1) 非安全模式。不采用信息安全管理和执行安全保护,设备运行一般应用时使用。该模式下,设备避开链路层的安全功能,可访问不含敏感信息的数据库。

(2) 业务层安全模式(也称为服务层安全模式)。蓝牙设备在 L2CAP 层建立信道之后采用信息安全管理并执行安全保护。这种安全机制建立在 L2CAP 及其之上的协议中,该模式可为多种应用提供不同的访问策略,可并行运行安全需求不同的应用。

(3) 链路层安全模式。在 LMP 层建立链路的同时就采用信息安全管理和执行安全保护,该安全机制建立在芯片中和 LMP 协议基础上。在该模式中,链路管理器(LM)在同一层面上对所有的应用强制执行安全措施。

安全模式 2 与安全模式 3 的本质区别在于：安全模式 2 下的蓝牙设备在信道建立以前启动安全过程,即它的安全过程在较高层协议进行;安全模式 3 下的蓝牙设备在信道建立以后启动安全过程,即它的安全过程在较低层协议进行。

蓝牙的安全信息单元包括：(1)蓝牙设备地址(BD_ADDR)：每个蓝牙设备唯一的 48 位的地址。(2)链路密钥：是身份认证和加密的重要参数,128 位。通过 E21 或 E22 算法产生。(3)加密密钥：用于加密的 8～128 位密钥,由 E3 算法产生。(4)随机数：由蓝牙设备产生的 128 位伪随机数。蓝牙安全机制的具体解释参见本章参考文献[1]。

7.1.2 ZigBee 安全简介

1. ZigBee 技术简介

2000 年 IEEE 成立了 IEEE 802.15.4 工作组,致力于开发一种可应用在固定便携或移动设备上的低成本低功耗及多结点的低速率无线个域网(LR-WPAN)技术标准,但该工作组只专注 MAC 层和物理层协议,要达到产品的互操作和兼容,还需要定义高层的规范。2001 年美国霍尼韦尔等公司发起成立了 ZigBee 联盟。ZigBee 联盟所主导的 ZigBee 标准定义了网络层、安全层、应用层和各种应用产品的资料或规范,并对其网络层协议和

应用编程接口(API)进行了标准化。2003年IEEE发布了IEEE Std 802.15.4—2003标准;2004年ZigBee1.0标准正式公布。许多世界著名硬件厂商都推出了支持ZigBee的无线产品,如Chipcon的CC2420、Freescale的MC13192、Atmel的AT86RF210等。

值得注意的是,ZigBee提供了高可靠性的安全服务,它的安全服务所提供的方法包括密码建立、密码传输、帧保护和设备管理。这些服务构成了一个模块用于实现ZigBee设备的各类安全策略。

2. ZigBee技术的特点

ZigBee技术不仅具有低成本、低功耗、低速率、低复杂度的特点,而且具有可靠性高、组网简单、灵活的优势。ZigBee技术和其他无线通信技术相比有如下特点。

(1) 低功耗:由于ZigBee的传输速率低,发射功率仅为1mW,而且采用了休眠模式,功耗低,因此ZigBee设备非常省电。据估算,ZigBee设备仅靠两节5号电池就可以维持长达6个月到2年的使用时间。

(2) 成本低:ZigBee模块的成本低廉,且ZigBee协议是免专利费的。

(3) 时延短:通信时延和从休眠状态激活的时延都非常短,典型的搜索设备时延为30ms,蓝牙需要3~10s,Wi-Fi需要3s;休眠激活的时延是15ms,活动设备信道接入的时延为15ms。因此ZigBee技术适用于对时延要求苛刻的无线控制(如工业控制场合等)应用。

(4) 低速率:ZigBee作在20~250Kbps的较低速率,分别提供250Kbps(2.4GHz)、40Kbps(915MHz)和20Kbps(868MHz)的数据吞吐率,满足低速率传输数据的应用需求。

(5) 近距离:传输范围一般介于10~100m之间,在增加发射功率后,亦可增加到1~3km。

(6) 网络容量大:一个星形结构的ZigBee网络最多可以容纳254个从设备和一个主设备,且网络组成灵活,一个区域内最多可以同时存在100个独立而且互相重叠覆盖的ZigBee网络。这一点与蓝牙相比优势明显。

正因为上述特点,Zigbee在无线传感器网络的组网结点中有大量的应用。

3. ZigBee安全架构

ZigBee协议栈的体系结构如图7.1所示。

IEEE 802.15.4—2003标准定义了最下面的两层:物理层(PHY)和MAC层。ZigBee联盟在此基础上建立了网络层(NWK层)和应用层(APL)框架。PHY层提供基本的物理无线通信能力;MAC层提供设备间的可靠性授权和单跳通信连接服务。NWK层提供用于构建不同网络拓扑结构的路由和多跳功能。应用层的框架包括了应用支持子层(APS)、ZigBee设备对象(ZDO)和由制造商制订的应用对象。ZDO负责所有设备的管理。APS提供ZigBee应用的基础。具体有三层安全机制:MAC、NWK和APL负责各自帧的安全传输。而且,APS子层提供建立和保持安全关系的服务;ZDO管理安全性策略和设备的安全性结构(部分名称的解释可查阅相关标准)。

4. 安全密钥

网络中ZigBee设备中的安全性是以一些“连接”密钥(Link Key)和一个“网络”密钥

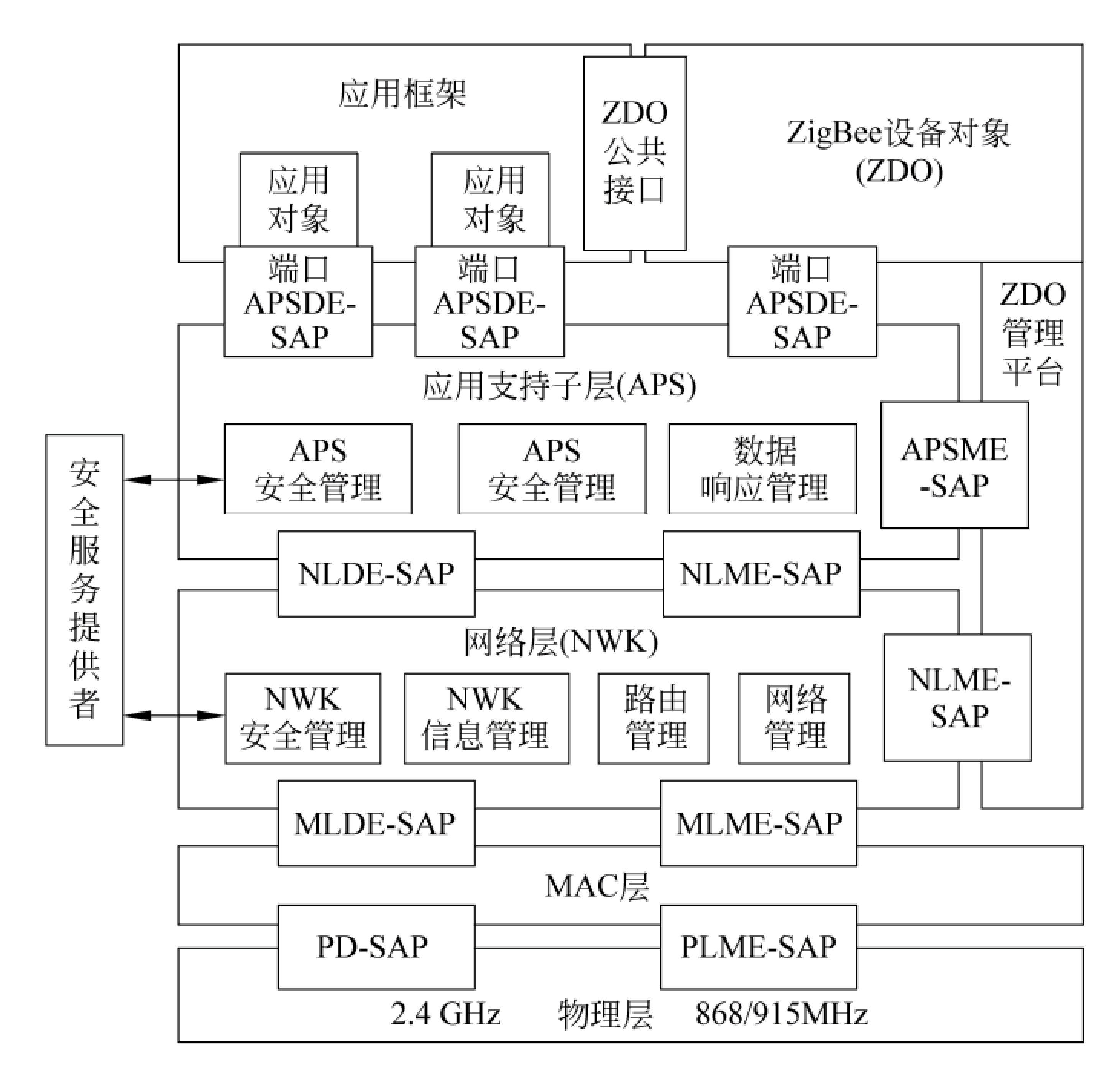


图 7.1　ZigBee 协议栈结构图

(Network Key)为基础的。应用层(APL)对等实体间的单播通信安全依靠由两个设备共享的一个 128 位连接密钥保证,而广播通信安全则依靠由网络中所有设备共享的一个 128 位网络密钥保证。接收方通常知道帧是被连接密钥保护还是网络密钥保护。

设备获得连接密钥可以通过密钥传输、密钥协商或者预安装等方式中的任意一种。网络密钥通过密钥传输或者预安装的方式获得。用于获取连接密钥的密钥协商技术基于主密钥(Master Key)。一个设备将通过密钥传输或者预安装方式获取一个主密钥来指定相应的连接密钥。设备间的安全性就是依靠这些密钥的安全初始化和安装来实现。

网络密钥可能被 ZigBee 的 MAC、NWK 和 APL 层使用,也就是说,相同的网络密钥和相关联的输入/输出帧计数器对所有层都是有效的。连接密钥和主密钥可能只被 APS 子层使用。连接密钥和主密钥只在 APL 层有效。

ZigBee 技术针对不同的应用,提供了不同的安全服务。这些服务分别作用在 MAC 层、NWK 层和 APL 层上,对数据加密和完整性保护是在 CCM 模式下执行 AES-128 加密算法。

5. NWK 层安全

当来自 NWK 层的帧需要保护,或者来自更高层的帧且网络层信息库(Network Information Base,NIB)中属性为 TRUE 时,ZigBee 使用帧保护机制。NIB 中的属性给出保护 NWK 帧的安全级别。上层通过建立网络密钥,决定使用哪个安全级别来保护 NWK 层。

通过多跳连接传送消息是 NWK 层的一个职责,NWK 层会广播路由请求信息并处理收到的路由回复消息。同时,路由请求消息会广播到其他设备,邻近设备则回复路由应

答消息。若连接密钥使用适当,NWK 层将使用连接密钥保护 NWK 帧的安全,若没有适当的连接密钥,为了保护信息,NWK 将使用活动的网络密钥保护 NWK 帧。帧的格式明确给出保护帧的密钥,因此接收方可以推断出处理帧的密钥。

6. APL 层安全

当来自 APL 层的帧需要安全保护时,APS 子层将会处理其安全性。APS 层的帧保护机制是基于连接密钥或网络密钥的。APS 层提供 ZDO 的密钥建立、密钥传输和设备管理等服务。

最后讨论两个容易混淆的问题:

① ZigBee 与 IEEE 802.15.4 的区别。ZigBee 建立在 IEEE 802.15.4 标准之上。IEEE 802.15.4 是 IEEE 确定的低速无线个域网(Personal Area Network)标准,这个标准定义了 PHY 层和 MAC 层。PHY 层规范主要是确定了 2.4GHz 以 250Kbps 的基准传输率。MAC 规范定义了同一区域工作的多个 802.15.4 无线电信号如何共享无线信道。支持几种架构,包括星型拓扑、树状拓扑、网状拓扑。仅仅定义 PHY 和 MAC 不足以保证不同的设备之间可以对话,于是便有了 ZigBee 联盟。ZigBee 从 802.15.4 标准入手,定义允许不同厂商设备相互对话的应用规范。ZigBee 联盟规范了网络层和应用层。协议中有作为 HUB 功能的协调器结点和基本结点。每个协调器结点可以连接多达 255 个结点,几个协调器结点可以构成一个网络。ZigBee 联盟还定义了安全层。

② ZigBee 正向 IP 迁移。ZigBee 架构与 IP 架构是不兼容的,这导致 ZigBee 网络与基于 IP 的服务与应用一起部署时,会产生严重的问题。基于 ZigBee 的网络与 IP 网络需要通过网关才能通信。网关需要完整的 ZigBee 协议栈和 IP 协议栈,这导致了很大的硬件和软件成本。因此,ZigBee 联盟在 2009 年宣布在最新的智能电网应用规范中,将向基于 IP 的架构迁移。

7.2　有线网络接入安全

物联网接入的特征在于网络的融合性,即网络接入方式可以是多种多样的。例如有线网络接入方法还包括 PSTN(Public Switched Telephone Network,公共交换电话网络)拨号接入、ADSL(Asymmetric Digital Subscriber Line,非对称数字用户线路)等公网接入方法。或者是专网接入(如校园网、中国教育科研网 CERNET 等)。在三网融合(电信网、有线电视网、计算机网络)的目标下,有线电视(Community Antenna Television,CATV)也可以连接互联网。还有各种独立网络布线商的高速宽带接入(如光纤到户 Fiber To The Home,FTTH,长城宽带等)。还有一种特殊的接入方式是电力线通信(Power Line Communication,PLC)接入方式,即通过电力线传送数据的方法。

由于物联网的一个基本应用就是改造传统的工业控制领域,如传感器和制动器的组网、远程控制、无线控制等。美国总统奥巴马 2009 年的演讲中多次提到的 M2M 的概念,非常强调机器间的通信。与之类似,美国国家自然科学基金委员会重点支持的学术领域 CPS 也是强调计算、通信、控制的融合,CPS 是一个以通信与计算为核心的集成的监控和

协调行动工程化物理系统。我国工业与信息化部在2009年提出的“两化融合”(即自动化和信息化的融合)的思想,也反映了控制网络通信的重要性。M2M和两化融合也被我国视为物联网4大技术的组成部分和应用领域之一。两化融合最基础的传统技术是基于短距离有线通信的现场总线的各种控制系统,如PLC,DCS(Distributed Control System),SCADA(Supervisory Control And Data Acquisition)等。物联网理念把信息技术融合到控制系统中,实现“高效、安全、节能、环保”的“管、控、营”一体化。

因此,本小节简介常常被人忽视的工业控制领域的有线网络安全。

7.2.1 现场总线简介

IEC(International Electrotechnical Commission,国际电工委员会)对现场总线(Fieldbus)的定义为:现场总线是一种应用于生产现场,在现场设备之间、现场设备和控制装置之间实行双向、串形、多结点的数字通信技术。它在生产现场、微机化测量控制设备之间实现双向串行多结点数字通信,也被称为开放式、数字化、多点通信的底层控制网络。在制造业、工业流程、交通、楼宇等方面的自动化系统中具有广泛的应用。现场总线不同于计算机网络,人们必须面对多种总线技术标准共存的现实。技术发展很大程度上受到市场规律、商业利益的制约;技术标准不仅是一个技术规范,也是一个商业利益的妥协产物。常见的现场总线有如下。

1. 基金会现场总线(Foundation Fieldbus,FF)

以美国Fisher-Rousemount公司为首联合了横河、ABB、西门子等80家公司制定的ISP协议,和以Honeywell公司为首的联合欧洲等地150余家公司制定的WorldFIP协议,于1994年9月合并而成。该总线在过程自动化领域得到了广泛的应用,具有良好的发展前景。

基金会现场总线采用国际标准化组织ISO的开放系统互联OSI的简化模型(1、2、7层),即物理层、数据链路层、应用层,另外增加了用户层。FF分低速H1和高速H2两种通信速率,前者传输速率为31.25Kbps,通信距离可达1900m,可支持总线供电和本质安全防爆环境。后者传输速率为1Mbps和2.5Mbps,通信距离为750m和500m,支持双绞线、光缆和无线发射,协议符号IEC1158—2标准。FF的物理媒介的传输信号采用曼彻斯特编码。

2. CAN(Controller Area Network,控制器局域网)

最早由德国BOSCH公司推出,它广泛用于离散控制领域,其总线规范已被ISO国际标准组织制定为国际标准,得到了Intel、Motorola、NEC等公司的支持。CAN协议分为二层:物理层和数据链路层。CAN的信号传输采用短帧结构,传输时间短,具有自动关闭功能,具有较强的抗干扰能力。CAN采用了非破坏性总线仲裁技术,通过设置优先级来避免冲突,通信距离最远可达10千米,通信速率最高可达40Mbps,网络结点数实际可达110个。

3. Lonworks

由美国Echelon公司推出,并由Motorola、Toshiba公司共同倡导。它采用ISO/OSI模型的全部7层通讯协议,采用面向对象的设计方法,通过网络变量把网络通信设计简化

为参数设置。支持双绞线、同轴电缆、光缆和红外线等多种通信介质，通讯速率从 300bps 至 1.5Mbps 不等，直接通信距离可达 2700m，被誉为通用控制网络。Lonworks 技术采用的 LonTalk 协议被封装到 Neuron 的芯片中，并得以实现。采用 Lonworks 技术和 Neuron 芯片的产品，被广泛应用在楼宇自动化、家庭自动化、保安系统、办公设备、交通运输、工业过程控制等行业。

4. DeviceNet

DeviceNet 是一种低成本的通信连接也是一种简单的网络解决方案，有着开放的网络标准。DeviceNet 基于 CAN 技术，传输率为 125Kbps 至 500Kbps，每个网络的最大结点数为 64 个，其通信模式为：生产者/客户（Producer/Consumer），采用多信道广播信息发送方式。位于 DeviceNet 网络上的设备可以自由连接或断开，不影响网上的其他设备，而且其设备的安装布线成本也较低。

5. PROFIBUS

PROFIBUS 是德国标准（DIN19245）和欧洲标准（EN50170）的现场总线标准。由 PROFIBUS-DP、PROFIBUS-FMS、PROFIBUS-PA 系列组成。DP 用于分散外设间高速数据传输，适用于加工自动化领域。FMS 适用于纺织、楼宇自动化、可编程控制器、低压开关等。PA 用于过程自动化的总线类型，服从 IEC1158-2 标准。PROFIBUS 支持主-从系统、纯主站系统、多主多从混合系统等几种传输方式。PROFIBUS 的传输速率为 9.6Kbps 至 12Mbps，最大传输距离在 9.6Kbps 下为 1200m，在 12Mbps 下为 200m，可采用中继器延长至 10km，传输介质为双绞线或者光缆，最多可挂接 127 个站点。

6. HART

HART（Highway Addressable Remote Transducer）最早由 Rosemount 公司开发。其特点是在现有模拟信号传输线上实现数字信号通信，属于模拟系统向数字系统转变的过渡产品。其通信模型采用物理层、数据链路层和应用层三层，支持点对点主从应答方式和多点广播方式。由于它采用模拟数字信号混合，难以开发通用的通信接口芯片。HART 能利用总线供电，可满足本质安全防爆的要求，并可用于由手持编程器与管理系统主机作为主设备的双主设备系统。

7. CC-Link

CC-Link（Control&Communication Link，控制与通信链路系统）由三菱电机为主导的多家公司于 1996 年 11 月推出，在亚洲占有较大份额。在其系统中，可以将控制和信息数据同是以 10Mbps 高速传送至现场网络，具有性能卓越、使用简单、应用广泛、节省成本等优点。其不仅解决了工业现场配线复杂的问题，同时具有优异的抗噪性能和兼容性。CC-Link 是一个以设备层为主的网络，同时也可覆盖较高层次的控制层和较低层次的传感层。2005 年 7 月 CC-Link 被中国国家标准委员会批准为中国国家标准指导性技术文件。

8. WorldFIP

WorldFIP 的北美部分与 ISP 合并为 FF 以后，WorldFIP 的欧洲部分仍保持独立，总部设在法国。其在欧洲市场占有重要地位，特别是在法国占有率大约为 60%。WorldFIP 的特点是具有单一的总线结构来适用不同的应用领域的需求，而且没有任何网关或网桥，

用软件的办法来解决高速和低速的衔接。WorldFIP 与 FFHSE 可以实现“透明连接”,并对 FF 的 H1 进行了技术拓展,如速率等。在与 IEC61158 第一类型的连接方面,WorldFIP 做得比较好。

9. INTERBUS

INTERBUS 是德国 Phoenix 公司推出的较早的现场总线,2000 年 2 月成为国际标准 IEC61158。INTERBUS 采用国际标准化组织 ISO 的开放系统互联 OSI 的简化模型(1、2、7 层),即物理层、数据链路层、应用层,具有强大的可靠性、可诊断性和易维护性。其采用数据环通信,具有低速度、高效率的特点,并严格保证了数据传输的同步性和周期性;该总线的实时性、抗干扰性和可维护性也非常出色。INTERBUS 广泛地应用到汽车、烟草、仓储、造纸、包装、食品等工业,成为国际现场总线的领先者。

此外较有影响的现场总线还有丹麦公司 Process-Data A/S 提出的 P-Net,该总线主要应用于农业、林业、水利、食品等行业;SwiftNet 现场总线主要使用在航空航天等领域。

目前现场总线通信安全的主流研究方向是通信的安全协议,该协议致力于开发安全检测措施以发现更多的传输错误。现有现场总线安全协议主要有 Profisafe,Interbus Safety,CANopen Safety,CCLink Safety,EtherCat Safety 等,其研究的重点均在传输错误的检测方法上。这些方法通常无法修复发生错误的信号,只能选择重传,且实时性有待提高。

7.2.2 工业控制系统安全简介

1. “震网”病毒事件

数据采集与监控(SCADA)、分布式控制系统(DCS)、过程控制系统(PCS)、可编程逻辑控制器(PLC,Process Control System)等工业控制系统广泛运用于工业、能源、交通、水利以及市政等领域,用于控制生产设备的运行。一旦工业控制系统中的软件漏洞被恶意代码所利用,将对工业生产运行和国家经济安全造成重大隐患。随着信息化与工业化深度融合以及物联网的快速发展,工业控制系统产品越来越多地采用通用协议、通用硬件和通用软件,以各种方式与互联网等公共网络相连接,于是,病毒、木马等威胁正在向工业控制系统扩散,工业控制系统信息安全问题日益突出。

2010 年发生的“震网”病毒事件就是一个典型的例子[8]。震网病毒又名 Stuxnet 病毒,它是第一个专门攻击现实世界中的工业基础设施的“蠕虫”病毒(能进行自我复制,通过网络传播),比如发电站和水厂。也被认为是世界上第一个网络“超级武器”,其目的可能是要攻击伊朗的布什尔核电站,它感染了全球超过 45000 个网络,其中伊朗遭到的攻击最为严重,60%的个人电脑感染了这种病毒。

Stuxnet 蠕虫针对的软件系统是西门子公司的 SCADA 系统 SIMATIC WinCC,该系统被广泛应用于钢铁、汽车、电力、运输、水利、化工、石油等工业领域,特别是国家基础设施工程中;该系统运行在 Windows NT 类型的平台上,常被部署在与外界隔离的专用局域网中。

Stuxnet 蠕虫利用了微软操作系统中至少 4 个漏洞,其中有 3 个全新的零日漏洞;伪造驱动程序的数字签名;通过一套完整的入侵和传播流程,突破工业专用局域网的物理限

制；利用 WinCC 系统的 2 个漏洞，对其开展破坏性攻击。通常情况下，蠕虫攻击意图在于传播范围的广阔性、攻击目标的普遍性。Stuxnet 蠕虫却与此截然相反，攻击的最终目标既不是开放主机，也不是大众家庭中的通用软件。而且攻击需要渗透到内网，且需要挖掘 Windows 操作系统的零日漏洞，表明攻击能力非同寻常。因而有猜测认为此次攻击很可能是一次带有政治意义的精心谋划的攻击。

工业控制网络，包括工业以太网以及现场总线控制系统早已在诸如电力、钢铁、化工等大型工业企业中应用多年，工控网络的核心大多是工控 PC，大多数同样基于 Windows-Intel 平台，工业以太网与民用以太网在技术上并无本质差异，现场总线技术更是将单片机/嵌入式系统应用到了每一个控制仪表上。以化工行业为例，针对工业控制网络的攻击可能破坏反应器的正常温度/压力测控，导致反应器超温/超压，最终就会导致冲料、起火甚至爆炸等灾难性事故，还可能造成次生灾害和人道主义灾难。

通过“震网”病毒事件，可以得到一些启示：

(1) 工业以太网和现场总线标准均为公开标准，熟悉工控系统的程序员开发具有针对性的恶意攻击代码并不存在很高的技术门槛。基于 Windows-Intel 平台的工控 PC 和工业以太网，可能遭到与民用或商用 PC 和网络中同样的攻击。因此，采取严格的网络隔离措施是必须的。

(2) DCS 和现场总线控制系统中测控软件的核心产品，特别是行业产品被少数公司所垄断，例如电力行业中常用的是西门子 SIMATIC WinCC。针对这种软件的攻击会从根本上破坏测控体系，Stuxnet 病毒的攻击目标正是 WinCC 系统。因此，在核心部门采用国产软件或者一定数量的异构异种软件，以及软件冗余措施是有益的。

(3) 基于 RS-485 总线以及光纤物理层的现场总线，例如 PROFIBUS 和 MODBUS，其安全性相对较好；但短程无线网络，特别是不使用 ZigBee 等通用短程无线协议（有一定的安全性），而使用自定义专用协议的短程无线通信测控仪表，安全性较差。

2. 工业控制系统的信息安全管理

工业与信息化部 2011 年 9 月下发了《关于加强工业控制系统信息安全管理的通知》[9]，要求各地区、各有关部门、有关国有大型企业充分认识工业控制系统信息安全的重要性和紧迫性，切实加强工业控制系统信息安全管理，以保障工业生产运行安全、国家经济安全和人民生命财产安全。重点加强核设施、钢铁、有色、化工、石油石化、电力、天然气、先进制造、水利枢纽、环境保护、铁路、城市轨道交通、民航、城市供水供气供热以及其他与国计民生紧密相关领域的工业控制系统信息安全管理，落实安全管理要求。通知从信息安全管理角度出发，提出了对工业控制系统的信息安全管理方面的注意事项，包括：

(1) 连接管理要求。

① 断开工业控制系统同公共网络之间的所有不必要连接。

② 对确实需要的连接，系统运营单位要逐一进行登记，采取设置防火墙、单向隔离等措施加以防护，并定期进行风险评估，不断完善防范措施。

③ 严格控制在工业控制系统和公共网络之间交叉使用移动存储介质以及便携式计算机。

(2) 组网管理要求。

① 工业控制系统组网时要同步规划、同步建设、同步运行安全防护措施。

② 采取虚拟专用网络(VPN)、线路冗余备份、数据加密等措施,加强对关键工业控制系统远程通信的保护。

③ 对无线组网采取严格的身份认证、安全监测等防护措施,防止经无线网络进行恶意入侵,尤其要防止通过侵入远程终端单元(RTU)进而控制部分或整个工业控制系统。

(3) 配置管理要求。

① 建立控制服务器等工业控制系统关键设备安全配置和审计制度。

② 严格账户管理,根据工作需要合理分类设置账户权限。

③ 严格口令管理,及时更改产品安装时的预设口令,杜绝弱口令、空口令。

④ 定期对账户、口令、端口、服务等进行检查,及时清理不必要的用户和管理员账户,停止无用的后台程序和进程,关闭无关的端口和服务。

(4) 设备选择与升级管理要求。

① 慎重选择工业控制系统设备,在供货合同中或以其他方式明确供应商应承担的信息安全责任和义务,确保产品安全可控。

② 加强对技术服务的信息安全管理,在安全得不到保证的情况下禁止采取远程在线服务。

③ 密切关注产品漏洞和补丁发布,严格软件升级、补丁安装管理,严防病毒、木马等恶意代码侵入。关键工业控制系统软件升级、补丁安装前要请专业技术机构进行安全评估和验证。

(5) 数据管理要求。

地理、矿产、原材料等国家基础数据以及其他重要敏感数据的采集、传输、存储、利用等,要采取访问权限控制、数据加密、安全审计、灾难备份等措施加以保护,切实维护个人权益、企业利益和国家信息资源安全。

(6) 应急管理要求。

制定工业控制系统信息安全应急预案,明确应急处置流程和临机处置权限,落实应急技术支撑队伍,根据实际情况采取必要的备机备件等容灾备份措施。

7.3 卫星通信接入安全

其实,卫星通信接入还包括卫星电视、遥感卫星等其他各类通信卫星技术,但民用卫星通信通常还是本节重点介绍的两个典型应用:CMMB手机电视和北斗卫星导航定位系统。

7.3.1 CMMB安全广播简介

CMMB(China Mobile Mulitimedia Broadcasting)技术标准(俗称手机电视)[10]是由中国广电总局组织研发,具有自主知识产权的移动多媒体广播电视标准,该标准适用于各

种7寸以下屏幕的移动便携终端,包括手机、GPS、MP4、数码相机等。CMMB具有全国统一标准、网络覆盖广、移动性好、节目丰富、终端方案成熟等特点。

从技术上说,手机电视主要分为两大类,分别源于广播网络和移动通信网络。第一类技术以欧洲的DVB-H,韩国的T-DMB、日本的ISDB-T、美国的MediaFlo和中国的CMMB为代表,实现上以地面广播网络为基础,与移动网络松耦合或者相对独立的组网;第二类技术以3GPP的MBMS和3GPP2的BCMCS为代表,以移动通信网络为基础,不能独立组网。与国外的同类技术如美、欧、日、韩等国相比,CMMB具有图像清晰流畅、组网灵活方便、内容丰富多彩的特点。

CMMB采用卫星和地面网络相结合的"天地一体、星网结合、统一标准、全国漫游"方式,实现全国范围移动多媒体广播电视信号的有效覆盖。CMMB利用大功率S波段卫星覆盖全国100%国土,利用S/U波段增补转发器覆盖卫星信号较弱区(利用UHF地面发射覆盖城市楼房密集区),利用无线移动通信网络构建回传通道,从而组成单向广播和双向交互相结合的移动多媒体广播网络。CMMB借助卫星和地面基站广播,解决了手机电视信号不流畅的问题。CMMB频段范围在470~798MHz,传播衰耗小,发射功率能够达到kW级别,有效室外覆盖范围在十几到40千米(覆盖范围大于移动通信基站)。到2011年12月,全国337个地级市和百强县实现优质覆盖,覆盖全国5亿以上的人口。CMMB也正在扩展海外业务,例如已经在塔吉克斯坦开通试播。

CMMB是广播技术,优势是覆盖广、相对成本低、多用户同时观看,不足之处是难以支持点播和双向互动的业务。3G的视频技术,优势在于交互性、点播甚至即时通信,但在实现大规模、广覆盖、多用户情况下的视频传输时是很不经济的。因此,将3G中的TD技术和CMMB技术结合起来,可加快移动电视在手机上的应用。

与传统的电视相比,CMMB除了传播音视频节目外,还可以利用自身的带宽优势提供各类数据业务,包括交通诱导、股市行情、电子杂志、生活咨询、推送式下载等。通过移动通信网络的回传,CMMB终端还可以实现互动、在线支付等功能。如此多样的应用必须要有安全的网络作为保障。如果没有安全广播,不法分子为了达到个人目的,就可能利用大功率的移动发射站在CMMB终端集中的区域进行非法广播,此时受干扰的终端收到的将是非法电台发出的毒害观众思想的电视节目,或者是联系电话及银行账号被篡改过的电视购物频道,此时受到伤害的不仅是消费者,运营商也会遭受到重大损失。因此,2009年1月广电总局颁布了《移动多媒体广播第10部分:安全广播》标准。

安全广播技术的原理就是:通过在移动多媒体广播信号中插入安全广播信息,使得移动多媒体终端具备鉴别多媒体广播业务合法性的能力。即当移动多媒体广播在传输过程被恶意替换、篡改时,终端可以及时停止非法业务的展现。

安全广播系统由前端子系统和终端模块构成,其中前端子系统实现安全广播信息的生成和发送,终端模块实现安全广播信息的接收和处理。安全广播前端子系统获得复用控制信息表和业务特征信息,并根据这些信息生成安全广播信息,以数据业务形式复用传输。复用子系统使用单独的复用子帧承载安全广播信息,为其分配业务标识号,经由广播信道发送。安全广播终端模块在终端接收移动多媒体广播业务内容时,根据业务标识号从传输帧中解复用获得安全广播信息,并对安全广播信息进行校验,根据校验结果确认广

播业务内容的合法性，进而允许或禁止业务展现，如图 7.2 所示。

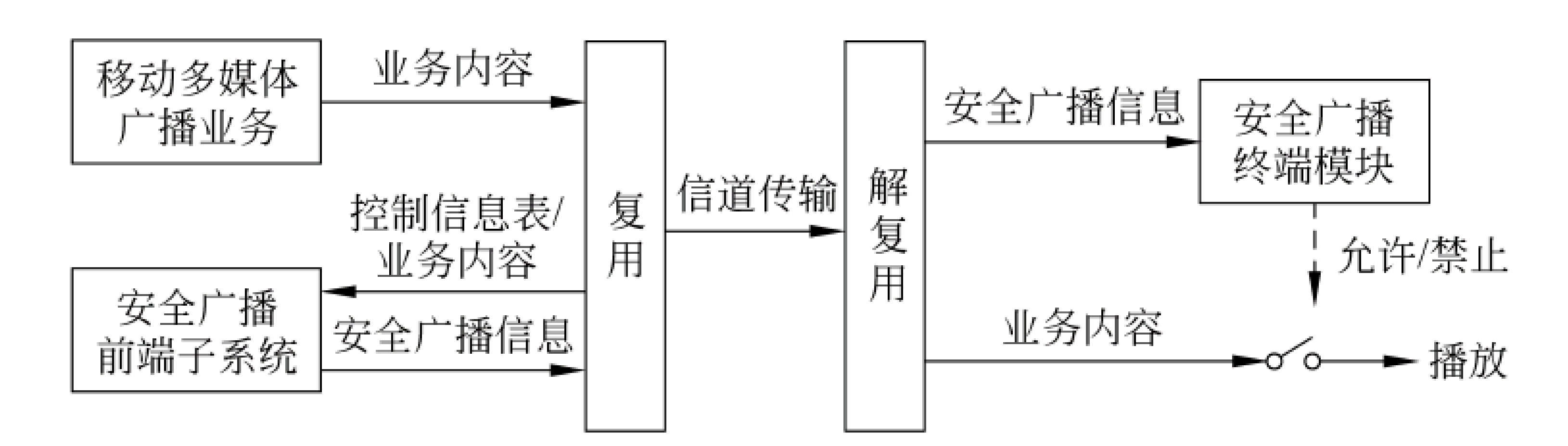


图 7.2 安全广播系统的基本组成

7.3.2 北斗卫星导航系统简介

北斗卫星导航系统(BeiDou COMPASS Navigation Satellite System)是中国正在实施的自主发展、独立运行的全球卫星导航系统。是除美国的全球定位系统 GPS、俄罗斯的 GLONASS 之后第三个成熟的卫星导航系统。可在全球范围内全天候、全天时为各类用户提供高精度、高可靠的定位、导航、授时服务，并兼具短报文通信能力。

该系统的建设目标是建成独立自主、开放兼容、技术先进、稳定可靠的覆盖全球的北斗卫星导航系统，促进卫星导航产业链形成，形成完善的国家卫星导航应用产业支撑、推广和保障体系，推动卫星导航在国民经济社会各行业的广泛应用。根据系统建设总体规划，2012 年左右，系统将首先具备覆盖亚太地区的定位、导航和授时以及短报文通信服务能力；2020 年左右，建成覆盖全球的北斗卫星导航系统。

北斗卫星导航系统由空间段、地面段和用户段三部分组成，空间段包括 5 颗静止轨道卫星和 30 颗非静止轨道卫星(卫星总数比 GPS 多出 11 颗)，地面段包括主控站、注入站和监测站等若干个地面站，用户段包括北斗用户终端以及与其他卫星导航系统兼容的终端。

北斗卫星导航系统的基本工作过程如图 7.3 所示。已经在轨使用的“北斗一号”系统采用的是主动式双向测距二维导航，首先由地面主控站向卫星Ⅰ和卫星Ⅱ同时发送询问信号，经卫星转发器向服务区内的用户广播。用户响应其中一颗卫星的询问信号，并同时向两颗卫星发送响应信号，经卫星转发回地面主控站。地面主控站接收并解释用户发来的信号，然后根据用户的申请服务内容进行相应的数据处理。对定位申请，主控站测出两个时间延迟：即从主控站发出询问信号，经某一颗卫星转发到达用户，用户发出定位响应信号，经同一颗卫星转发回主控站的延迟；和从主控站发出询问信号，经上述同一卫星到达用户，用户发出响应信号，经另一颗卫星转发回主控站的延迟。由于主控站和两颗卫星的位置均是已知的，因此，由上面两个延迟量可以算出用户到第一颗卫星的距离，以及用户到两颗卫星距离之和，从而知道用户处于一个以第一颗卫星为球心的一个

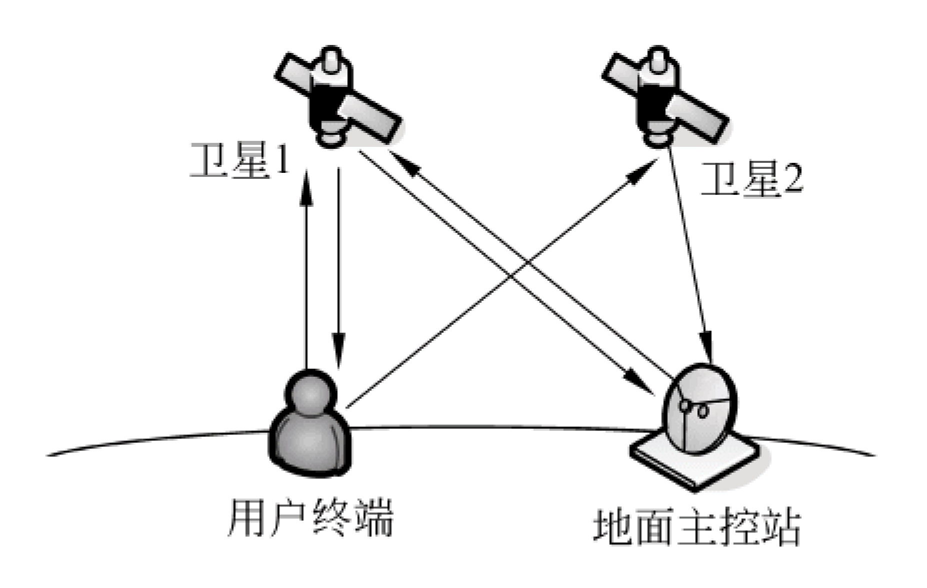


图 7.3 北斗卫星系统的基本工作过程

球面，和以两颗卫星为焦点的椭球面之间的交线上。另外主控站从存储在计算机内的数字化地形图查寻到用户高程值，又可知道用户处于某一与地球基准椭球面平行的椭球面上。从而中心控制系统可最终计算出用户所在点的三维坐标，这个坐标经加密由出站信号发送给用户。

北斗系统的 4 大功能如下。

(1) 短报文通信。北斗系统用户终端具有双向报文通信功能，用户可以一次传送 40～60 个汉字的短报文信息。这一功能是 GPS 所不具备的。

(2) 精密授时。北斗系统具有精密授时功能，可向用户提供 20～100ns 时间同步精度。

(3) 定位精度。水平精度 100 米，设立标校站之后为 20 米。工作频率为 2491.75MHz。

(4) 系统容纳的最大用户数为每小时 540000 户。

北斗系统在军事上有及其重要的用途，如航海定位、武器精确制导等。在民用方面包括个人位置服务、气象、道路交通管理、铁路运输、海运和水运、航空运输、应急救援等。

目前北斗系统的安全性的重心还是放在军事应用上，其具体的安全机制均处于保密状态，不为外界所知。随着其在民用网络的应用不断增加，安全机制的设计与研究也会逐渐提上日程。北斗系统应向终端用户提供安全、有效和可靠的数据服务。具体需要满足以下方面：

(1) 服务方可以认证终端用户的身份，保证未授权用户无法窃取数据服务，即数据的授权访问服务。

(2) 服务方向终端用户提供有效和可靠的加密数据，只有终端用户可以阅读和使用这个加密数据，未授权用户无法阅读和使用数据，即数据的保密通信服务。

(3) 终端用户可以认证加密数据的来源可靠，进而可以放心使用这个加密数据，即数据的来源认证服务。

参考文献[13]设计了一个数据安全方案，为终端用户提供授权访问、保密通信及数据来源认证等数据安全服务，使终端用户更加安全地使用北斗卫星导航系统。

研究与思考

[1] 利用参考文献[3]，开发一个实际的 ZigBee 应用。
[2] 基于 OMS 平台，如何利用 CMMB 进行手机应用开发。
[3] 利用北斗卫星导航系统开发一个智能手机上基于位置的服务应用程序。
[4] 为北斗卫星导航系统的民用服务设计一套安全机制。

进一步阅读建议

[1] N. Sastry and D. Wagner, Security Considerations for IEEE 802.15.4 Networks[C], In Proc. of WiSE'04, October 1, 2004, Philadelphia, Pennsylvania, USA.
[2] K. Haataja, P. Toivanen, Two practical man-in-the-middle attacks on Bluetooth secure simple

pairing and countermeasures [J], IEEE Transactions on Wireless Communications, 9(1): 384-392, January 2010.

本章参考文献

[1] Bluetooth SIG, Specification of the Bluetooth System Version 1.1[S]. 2001.
[2] K. Scarfone and J. Padgette, Guide to Bluetooth Security[S]. National Institute of Standards and Technology, NIST, USA.
[3] 李文仲,段朝玉. ZigBee 无线网络技术入门与实战[M]. 北京:北京航空航天大学出版社,2007.
[4] IEEE Computer Society. Standard for part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specification for low-rate wireless personal area networks (LR-WPANs) [S]. IEEE Std 802.15.4, 2003.
[5] 周洪波. 物联网:技术、应用、标准和商业模式[M]. 第 2 版. 北京:电子工业出版社,2011 年 7 月.
[6] 岳东峰,于东,高甜容,尹震宇,王品. 基于数控双环现场总线的安全通信方法[J]. 计算机集成制造系统. 17(5),2011 年 5 月,1032-1039.
[7] 李正军. 现场总线及其应用技术[M]. 北京:机械工业出版社,2006 年 1 月.
[8] 百度百科,震网病毒[OL],http://baike.baidu.com/view/4416726.htm.
[9] 工业与信息化部,关于加强工业控制系统信息安全管理的通知[OL]. http://www.miit.gov.cn/n11293472/n11293832/n12843926/n13917012/14294613.html.
[10] 中广传播集团有限公司[OL]. http://www.cbc.cn/.
[11] 安全广播在 CMMB 终端的应用与实现[OL]. http://www.21ic.com/app/ce/200905/43063.htm.
[12] 北斗卫星导航系统[OL]. http://www.beidou.gov.cn/.
[13] 程翔,陈恭亮,李建华,龚洁中. 基于北斗卫星导航系统的数据安全应用[J]. 信息安全与通信保密,2011 年 6 期.
[14] 刘云浩. 物联网导论[M]. 北京:科学出版社,2011.

第 8 章　物联网核心网安全——6LoWPAN 和 RPL 的安全性

前面的章节介绍了接入网的安全，本章介绍核心承载网络的安全。全 IP 网络是一个趋势，本章重点介绍核心网中最重要的最常见的是 IP 网络，这可归结到有线网络的安全，参考有线网络的安全解决方案。核心网中最重要的最常见的是 IP 网络，这可归结到有线网络的安全，参考有线网络的安全解决方案。有线网络的安全研究已经有比较长的时间，有比较成熟的研究成果和商业产品，如防火墙、入侵检测系统、入侵保护系统、虚拟专用网(Virtual Private Network，VPN)、网络隔离系统等。本章主要从分层协议的角度，从原理上介绍一下 IP 层和 TCP 层的安全增强机制。

在 2011 年 2 月，国际互联网协会 ICANN 官方宣布全球最后一批 IP 地址分配完毕，这标注着 IPv6 时代即将到来。本章的介绍主要针对 IPv6 协议(虽然介绍的机制也可以在 IPv4 上实现)。对于核心承载网而言，IPv6 还提供了 IPv4 更高的安全性。总之，IPv6 相对于 IPv4 而言，有很多优点：具有 128 位超大地址空间；支持更多的安全性；配置简单；提供认证和保密功能性；允许扩充；支持资源分配等。

除了核心网外，感知层也可能需要配置 IPv6 协议。例如对于使用智能物体(如 IPSO 联盟)架构的物联网来说，终端结点需要配置 IPv6 地址来提供对终端结点的广域网路由和寻址能力，例如基于 Internet 的传感器网络，传感器 Web 应用等(对于 EPCglobal 架构的物联网，终端的 IPv6 地址配置不是必须的)。对于存在网关的传感器网络而言，终端结点也可以不配备的 IPv6，只需要在网关上配备 IPv6 即可，再通过网关和结点间的 IEEE 802.15.4 协议相互访问。对于基于 M2M 架构的大规模传感器物联网而言，终端结点也可不必配备 IPv6，只要移动运营商那里具有结点的唯一标识(例如 SIM 卡的 IMSI 或者 USIM)即可寻址和访问到结点。

8.1　核心 IP 骨干网的安全

安全机制可以处在协议栈的不同层次，通常密钥协商和认证协议在应用层定义，而保密性和完整性可在不同的层次完成，见表 8.1。这里的认证对象是指是消息鉴别还是设备认证、用户认证等。

表 8.1　分层安全协议

所处层次	安全协议	应用对象	保密性	完整性	认证对象
应用层	WS-Security	文档	是	是	数据
	PGP	E-mail	是	是	消息
	S/MIME		是	是	

续表

<table>
<tr><th>所处层次</th><th>安全协议</th><th>应用对象</th><th>保密性</th><th>完整性</th><th>认证对象</th></tr>
<tr><td rowspan="2">传输层</td><td>SSH</td><td rowspan="3">客户端到服务器</td><td>是</td><td>否</td><td>用户</td></tr>
<tr><td>SSL/TLS</td><td>是</td><td>是</td><td rowspan="2">服务器</td></tr>
<tr><td>网络层</td><td>IPSec</td><td>是</td><td>是</td></tr>
<tr><td rowspan="2">链路层</td><td>WEP/WPA/802.1X
GSM/3G/LTE</td><td>主机到主机</td><td>是</td><td>是</td><td>主机</td></tr>
<tr><td>IEEE 802.15.4，
蓝牙</td><td>无线访问</td><td>是</td><td>是</td><td>设备</td></tr>
</table>

前面章节对链路层的安全机制介绍较多，下面重点介绍 IPSec 和 TLS/SSL 安全机制。

8.1.1 IPSec

从 1995 年开始，IETF 着手研究制定了一套 IP 安全（IP Security，IPSec）协议用于保护 IP 通信的安全。IPSec 将密码算法设立在网络层，它是构造 VPN 的主要工具。IETF 的 IPSec 工作组定义了 12 个 RFC，定义了体系、密钥管理、基本协议等，因此，IPSec 是一种协议套件。

IPSec 由三个基本组成部分：认证报头（Authentication Header，AH）、封装安全负载（Encapsulating Security Payload，ESP）和 Internet 密钥交换（Internet Key Exchange，IKE）协议。IPSec 提供的安全服务包括：数据起源地验证、无连接数据的完整性验证、数据内容的机密性、抗重播保护和有限的数据流机密性保证等。

IPSec 具有两种工作模式，即传输模式和隧道模式。传输模式不保护 IP 头，只保护 IP 包中的来自传输层的数据包（IP 层载荷）。通常用于从主机到主机的数据保护场景。隧道模式保护整个 IP 包（包含原 IP 头），因而要加一个新的 IP 头，通常使用在从主机到路由器和从路由器到主机的路由器上，也就是说当发送者和接收者都不是主机的时候，才使用隧道模式。

1. 认证报头

AH 可以对数据包提供完整性验证、数据源认证、选择性抗重播服务。AH 通常采用 HMAC-MD5-96 或 HMAC-SHA1-96 算法获得数据完整性的验证值。该验证算法需要用到发送与接收方共享的密码。AH 中设有序列号域，使目的主机可以辨别哪些是有效的数据包，拒收重复传播数据包，从而防御重放攻击。

在介绍 AH 的格式之前，需要明确一个概念：安全关联（Security Association，SA）。它是一种通信双方达成的一套安全共识，包括 3 个参数：

（1）安全参数索引（Security Parameter Index，SPI）。SPI 是安全算法和参数的集合的一个编号，使得通信两端主机运行的 IPSec 根据编号可以知道使用哪个算法和哪个参数进行秘密通信。

(2) 目标 IP 地址。它用于标明该 SA 是针对哪个终端主机。

(3) 安全协议标识符。它用于标明该 SA 采用的是 AH 还是 ESP。

图 8.1 给出了 AH 的格式。各域的解释如下。

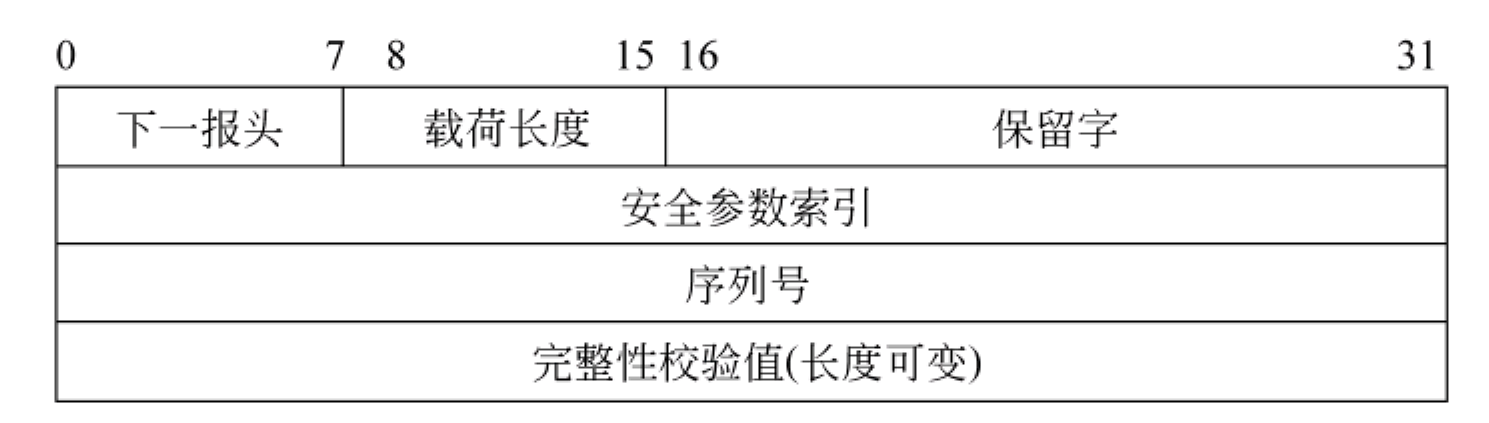


图 8.1　AH 的格式

(1) 下一报头：该域的长度是 8 位。表明紧跟在认证头后面的下一个载荷(payload)的类型。

(2) 载荷长度：该域的长度是 8 位。指明 AH 的长度，但并不包括第一个 8 字节。例如若完整性校验值为 96 位的 HMAC，则这里的值为 96/32+1=4。

(3) 保留字：该域的长度是 16 位，保留将来使用，目前为 0。

(4) 安全参数索引(SPI)：该域的长度是 32 位。它把目的地址和安全协议结合起来，唯一的指明了该数据报的安全关联。

(5) 序列号：该域的值是无符号 32 位整数，表示一个递增的计数器值。

(6) 完整性校验值，又叫认证数据，该域包含了对该数据报的完整性校验值(ICV，Integrity Check Value)。该域的长度可变，但必须是 32 位的整数倍，如用 HMAC-SHA-1 计算后输出前 96 位。

AH 头的位置依赖于 AH 的操作模式。AH 有两种操作模式：传输模式和隧道模式。

传输模式适合于不经过网关、路由器等中心设备的两台主机之间，应用范围比较小；而隧道模式能适合几乎所有的场合。AH 传输模式中，AH 被插在 IP 头之后传输层协议之前，或者所有其他 IPSec 协议头之前。

AH 隧道模式既可以由主机，也可以由安全网关使用。当 AH 被实现为安全网关来保护传输流量时，必须使用隧道模式。在隧道模式中，AH 插在原始的 IP 头之前，另外生成一个新的 IP 头放在 AH 之前。图 8.2 是加入 AH 后的报文示意图。



图 8.2　AH 加入后的 IP 包

2. 封装安全载荷

ESP 提供数据加密功能、数据认证、完整性和抗重播功能以及使用安全性网关提供有限的数据流机密性。ESP 比 AH 多一个功能，即保密性。封装安全载荷头的格式如图 8.3 所示。各域的解释如下。

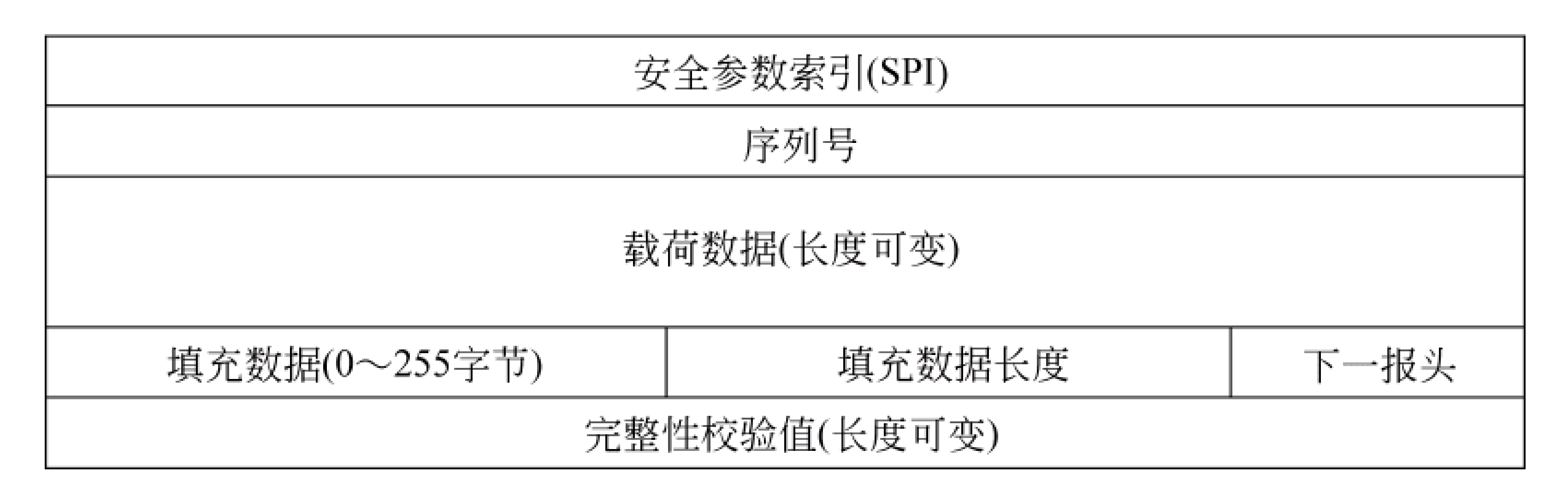


图 8.3 ESP 的格式

(1) 安全参数索引(SPI)：SPI 是一个任意的 32 位值，它唯一标识用于该数据报的 SA。

(2) 序列号：这是一个无符号的 32 位字段，表明一个递增的计数器值。

(3) 载荷数据：它包含下一个头字段描述的数据。有效载荷数据字段是强制性的，它的长度是字节的整数倍。

(4) 填充数据：为了使载荷在正确的位置结束，可以使用填充数据。发送者可以填充 0～255 字节的填充数据。

(5) 填充数据长度：这个字段指示放在该字段前面的填充字节的数量。有效值范围为 0～255，0 表示没有填充字节。

(6) 下一报头：这个字段标识包含在载荷数据字段中的数据的类型，也就是说，表明是 IPv6 的扩展头部还是上层协议标志符。

(7) 完整性校验值，又叫认证数据：它包含了对 ESP 减去认证数据部分后剩余数据计算得到的 ICV。

同样的，根据 ESP 的位置不同，ESP 也有两种操作模式：传输模式与隧道模式。

传输模式用于加密和认证 IP 携带的数据，在 IPv6 中，ESP 被用作端到端载荷，即不被中间路由器校验和处理。ESP 头出现在 IPv6 基本头、跳、路由和分段扩展头之后，目的可选扩展头可根据需要出现在 ESP 头之前或之后。可选扩展头在 ESP 头之后，加密包括整个传输段、ESP 尾和目的可选扩展头。ESP 头和所有密文将都被认证。

在隧道模式下，ESP 被插在原始 IP 头之前，并且生成一个新的 IP 头并将其插在 ESP 之前。隧道模式的认证和加密服务要对整个内部 IP 头进行认证和加密；而外部 IP 头既未被认证也未被加密。但是 ESP 隧道模式认证和加密服务所提供的安全性强于 ESP 传输模式。ESP 隧道模式的保密服务，特别在安全网关上实现时，可以提供数据流保密服务，因为包含 IP 数据包源地址的内部 IP 头被加密了。图 8.4 是加入 ESP 后的报文示意图。

加密算法和 HMAC 数据源和数据完整性验证都要求通信双方拥有共享的密钥。这

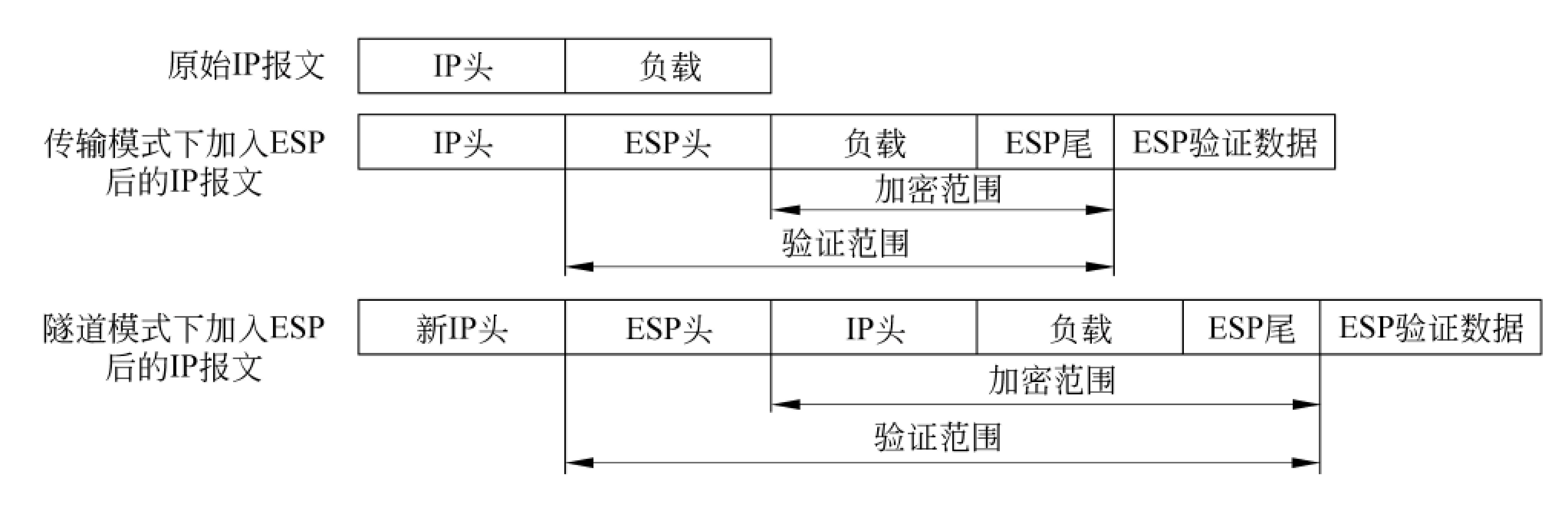


图 8.4　ESP 加入后的 IP 包

个过程可以通过密钥交换协议完成，或者由系统管理员用人工方式实现分配密钥。IPSec 使用了 IKE 协议，由 Oakley 协议、SKEME（安全密钥交换机制）和 ISAKMP（Internet 安全关联和密钥管理协议）组成。IKE 协议可用来动态地建立 SA，从而简化了 IPSec 的配置和维护工作。Oakley 协议来自于 Diffie-Hellman 密钥交换协议和身份认证协议，但没有规定协议的格式。ISAKMP 协议明确规定了协议的格式（即规定了包头中的域内容及其长度），但没有规定密钥交换算法。具体内容参见相关文档。

8.1.2　SSL/TLS

传输层安全协议通常指的是套接层安全协议 SSL 和传输层安全 TLS 协议两个协议。SSL 是美国 Netscape 公司于 1994 年设计的，为应用层数据提供安全服务和压缩服务。SSL 虽通常是从 HTTP 接收数据，但 SSL 其实可以从任何应用层协议接收数据。IETF 于 1999 年将 SSL 的第 3 版进行了标准化，确定为传输层标准安全协议 TLS。TLS 和 SSL 第 3 版只有微小的差别，故人们通常把它们一起表示为 SSL/TLS。另外，在无线环境下，由于手机及手持设备的处理和存储能力有限，原 WAP 论坛在 TLS 的基础上做了简化，提出了 WTLS 协议（Wireless Transport Layer Security），以适应无线网络的特殊环境。

SSL 由两部分组成，第一部分称为 SSL 记录协议，置于传输协议之上，第二部分由 SSL 握手协议、SSL 密钥更新协议和 SSL 提醒协议组成，置于 SSL 记录协议之上和应用程序如 HTTP 之下。图 8.5 显示了 SSL 协议在应用层和传输层之间的位置。

HTTP		
SSL握手协议	SSL密钥更新协议	SSL提醒协议
SSL记录协议		
TCP		
IP		

图 8.5　SSL 协议结构

1. SSL 握手协议

SSL 握手协议用于给通信双方约定使用哪个加密算法、哪个数据压缩算法以及哪些

参数。在算法确定了加密算法、压缩算法和参数以后，SSL 记录协议将接管双方的通信，包括将大数据分割成块、压缩每个数据块、给每个压缩后的数据块签名、在数据块前加上记录协议包头并传送给对方。SSL 密码更换协议允许通信双方在一个会话阶段中更换算法或参数。SSL 提醒协议是管理协议，用于通知对方在通信中出现的问题以及异常情况。

SSL 握手协议是 SSL 各协议中最复杂的协议，它提供客户和服务器认证并允许双方协商使用哪一组密码算法，交换加密密钥等。它分 4 个阶段，如图 8.6 所示。

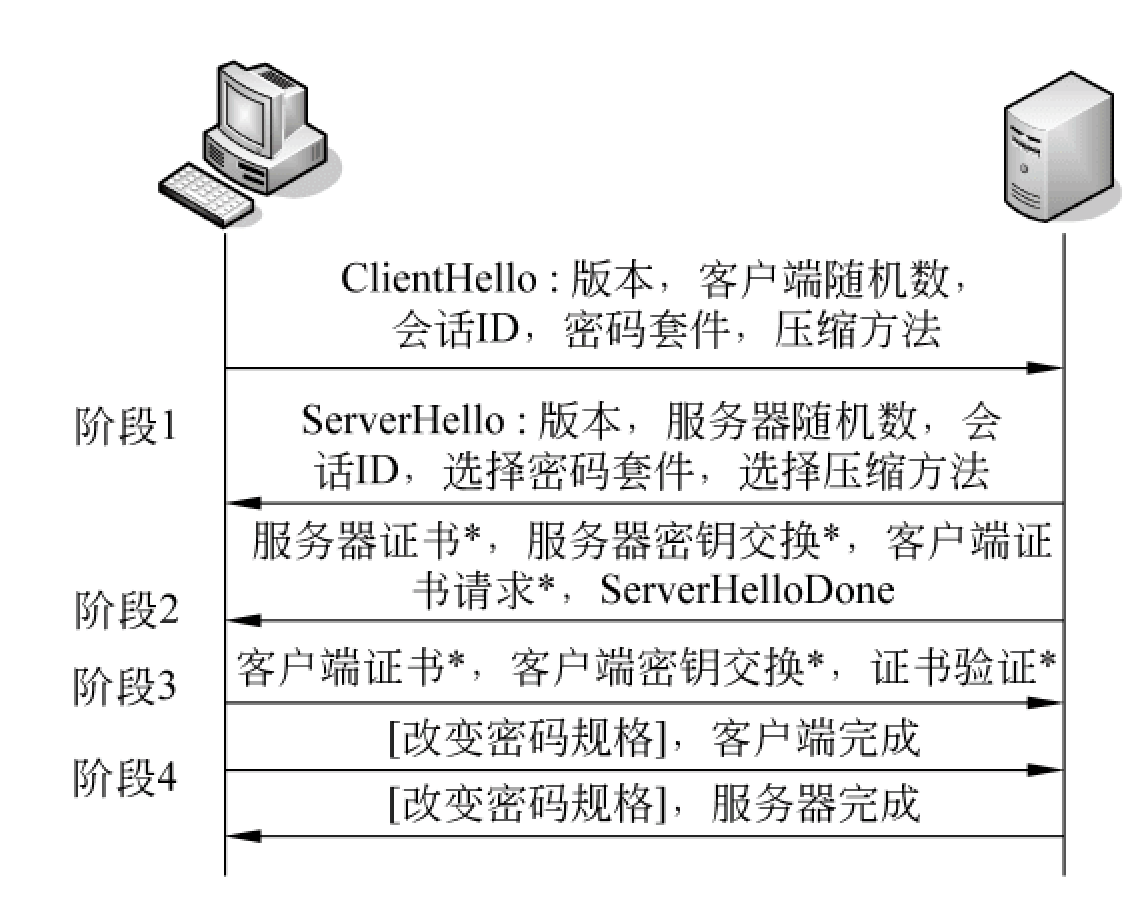


图 8.6 SSL 握手协议(带 * 号是可选的，[]不是 TLS 消息)

(1) 第 1 阶段：协商确定双方将要使用的密码算法。

这一阶段的目的是客户端和服务器各自宣布自己的安全能力，从而双方可以建立共同支持的安全参数。客户端首先向服务器发送问候信息，包括：客户端主机安装的 SSL 最高版本号，客户端伪随机数生成器秘密产生的一个随机串 r_c 防止重放攻击，会话标识，密码算法组，压缩算法(ZIP、PKZIP 等)。其中密码算法组是指客户端主机支持的所有公钥密码算法、对称加密算法和 Hash 函数算法。按优先顺序排列，排在第一位的算法客户主机最希望使用的算法。例如，客户的 3 种算法分别为：

- 公钥密码算法：RSA、ECC、Diffie-Hellman；
- 对称密码算法：AES-128、3DES/3、RC5；
- Hash 函数算法：SHA-512，SHA-1，MD5。

然后，服务器向客户端回送问候信息。包括：服务器主机安全的 SSL 最高版本号，服务器伪随机数生成器秘密产生的随机串 r_s，会话标识，密码算法组，例如 RSA、3DES/3、SHA-1，压缩算法。

(2) 第 2 阶段：对服务器的认证和密钥交换。

服务器程序向客户程序发送如下信息：

① 服务器的公钥证书。包含 X.509 类型的证书列表，如果密钥交换算法是匿名 Diffie-Hellman，就不需要证书。

② 服务器端的密钥交换信息。包括对预备主密钥的分配。如果密钥交换方法是

RSA 或者固定 Diffie-Hellman，就不需要这个信息。

③ 询问客户端的公钥证书。向客户端请求第 3 阶段的证书。如果客户使用的是匿名 Diffie-Hellman，服务器就不向客户端请求证书。

④ 完成服务器问候。该信息用 ServerHelloDone 表示，表示阶段 2 结束，阶段 3 开始。

（3）第 3 阶段：对客户端的认证和密钥交换。

客户程序向服务器程序发送如下信息：

① 客户公钥证书。和第 2 阶段第①步信息格式相同，但内容不同，它包含证明客户的证书链。只有在第 2 阶段第③步请求了客户端的证书，才发送这个信息。如果有证书请求，但客户没有可发送的证书，它就发送一个 SSL 提醒信息（携带一个没有证书的警告）。服务器也许会继续这个会话，也可能会决定终止。

② 客户端密钥交换信息。用于产生双方将使用的主密钥，包含对预备主密钥的贡献。信息的内容基于所用的密钥交换算法。如果密钥交换算法是 RSA，客户就创建完整的预备主密钥并用服务器 RSA 公钥进行加密。如果是匿名 Diffie-Hellman 或暂时 Diffie-Hellman，客户就发送 Diffie-Hellman 半密钥等。

③ 证书验证。如果客户发送了一个证书，宣布它拥有证书中的公钥，就需要证实它知道相关的私钥。这对于阻止一个发送了证书并声称该证书来自客户的假冒者是必需的。通过创建一个信息并用私钥对该信息进行签名，可证明它拥有私钥。例如客户用私钥对前面发送的明文的 Hash 值进行签名。

假设服务器在第 1 阶段选取了 RSA 作为密钥交换手段，则客户程序用如下方法产生密钥交换信息。客户程序验证服务器公钥证书的服务器公钥，然后用伪随机数生成器产生一个 48 字节长的比特字符串 s_{pm}，称为前主密钥。然后用服务器公钥加密 s_{pm}，将密文作为密钥交换信息传给服务器。这时，客户端和服务器端均拥有 r_c，r_s，s_{pm}，且 s_{pm} 是仅仅被客户和服务器所拥有。此后，双方计算主密钥 s_m：

$$
\begin{aligned}
s_m = & H_1(s_{pm} \parallel H_2(\text{‘A’} \parallel s_{pm} \parallel r_c \parallel r_s)) \parallel \\
& H_1(s_{pm} \parallel H_2(\text{‘BB’} \parallel s_{pm} \parallel r_c \parallel r_s)) \parallel \\
& H_1(s_{pm} \parallel H_2(\text{‘CCC’} \parallel s_{pm} \parallel r_c \parallel r_s))
\end{aligned}
$$

其中，H_1 和 H_2 是 Hash 函数（SSL 用 MD5 作为 H_1 的默认 Hash 函数，用 SHA-1 作为 H_1 的默认 Hash 函数），‘A’，‘BB’，‘CCC’分别表示 A、BB、CCC 的 ASCII 码。

（4）第 4 阶段：结束。

双方互送结束信息完成握手协议，并确认双方计算的主密钥相同。为达到此目的，结束信息将包含双方计算的主密钥的 Hash 值。

握手协议完成后，双方用产生主密钥 s_m 的方法，用 s_m 取代 s_{pm}，并根据双方商定的密码算法，产生一个足够长的密钥块 K_b 如下：

$$
\begin{aligned}
K_b = & H_1(s_m \parallel H_2(\text{‘A’} \parallel s_m \parallel r_c \parallel r_s)) \parallel \\
& H_1(s_m \parallel H_2(\text{‘BB’} \parallel s_m \parallel r_c \parallel r_s)) \parallel \\
& H_1(s_m \parallel H_2(\text{‘CCC’} \parallel s_m \parallel r_c \parallel r_s)) \\
& \cdots
\end{aligned}
$$

然后 SSL 将 K_b 分割成 6 段，每一段自成一个密钥。这 6 个密钥分成如下两组：第 1 组为 (K_{c1}, K_{c2}, K_{c3})；第 2 组为 (K_{s1}, K_{s2}, K_{s3})。每组 3 个密钥，即

$$K_b = K_{c1} \| K_{c2} \| K_{c3} \| K_{s1} \| K_{s2} \| K_{s3} \| Z$$

其中 Z 是剩余的字符串。

第 1 组密钥用于客户到服务器的通信，记为：

$$(K_{c1}, K_{c2}, K_{c3}) = (K_{cHMAC}, K_{cE}, IV_c)$$

分别为认证密钥，加密密钥，和初始向量。

第 2 组用于服务器到客户的通信，记为：

$$(K_{s1}, K_{s2}, K_{s3}) = (K_{sHMAC}, K_{sE}, IV_s)$$

作用和第 1 组类似。

此后，客户和服务器将转用 SSL 记录协议进行后续的通信。

2. SSL 记录协议

执行握手协议之后，客户和服务器双方统一了密码算法、算法参数、密钥及压缩算法。SSL 记录协议便可使用这些算法、参数和密钥对数据进行保密和认证处理。令 M 为客户希望传送给服务器的数据。客户端 SSL 记录协议首先将 M 分成若干长度不超过 2^{14} 字节的分段：$M_1, M_2, \cdots, M_k$。令 CX、H 和 E 分别为客户端和服务器双方在 SSL 握手协议中选定的压缩函数、HMAC 算法和加密算法。客户端 SSL 记录协议按如下步骤先将每段 M_i 进行压缩、认证和加密处理，然后将其发送给服务器，$i=1,2,\cdots,k$，如图 8.7 所示。

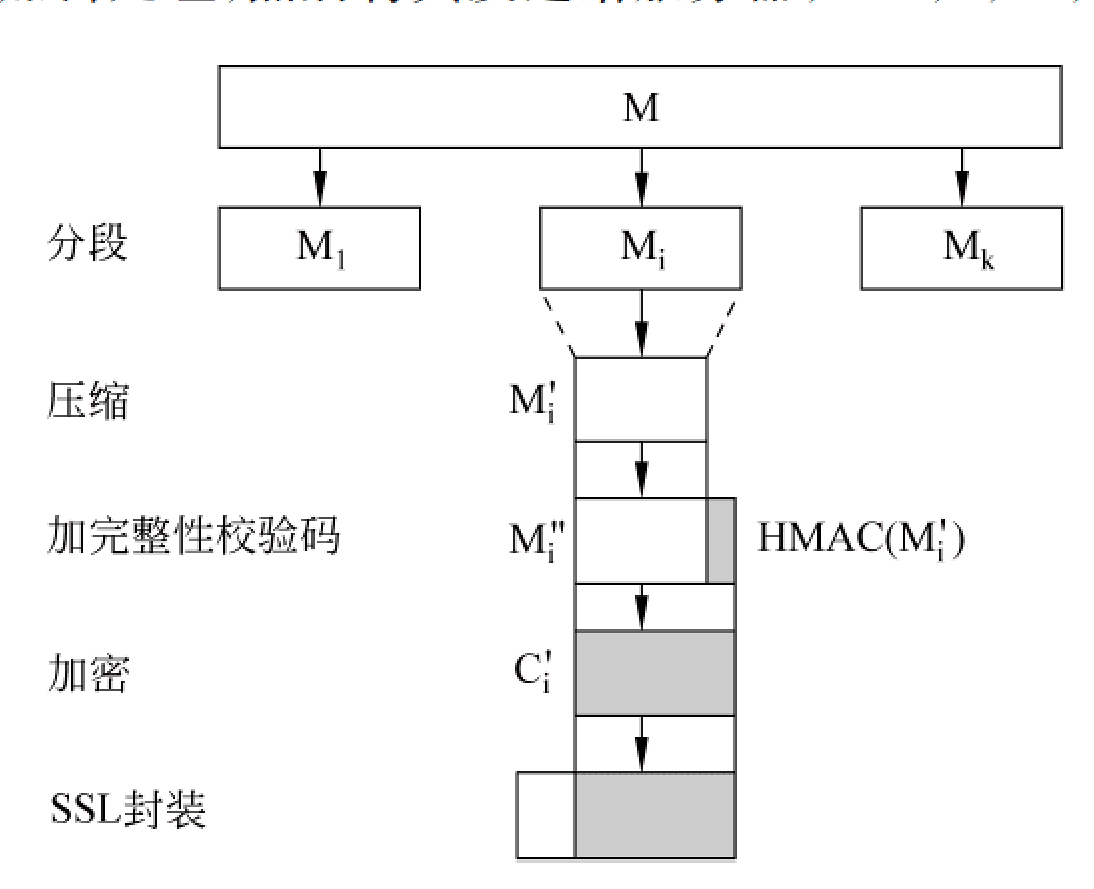


图 8.7　SSL 记录协议示意图

① 将 M_i 压缩得 $M_i' = CX(M_i)$。

② 将 M_i' 进行认证得 $M_i'' = M_i' \| H_{K_{cHMAC}}(M_i')$。

③ 将 M_i'' 加密得 $C_i = E_{K_{cE}}(M_i'')$。

④ 将 C_i 封装得 $P_i = [\text{SSL 记录协议包头}] \| C_i$。

⑤ 将 P_i 发送给服务器。

服务器收到客户送来的 SSL 记录协议包后，首先将 C_i 解密得 $M_i' \| H_{K_{cHMAC}}(M_i')$，验证 HMAC，然后将 M_i' 解压还原成 M_i。同理，从服务器发送给客户的数据也按上述方式处理。双方间的通信保密性和完整性由此得到保护。

8.2 6LoWPAN 适配层的安全

为了让 IPv6 协议在 IEEE 802.15.4 协议之上工作，导致了 6LoWPAN 适配层的提出。这一解决方法正在被 IPSO 联盟所推广，是 IPSO 提出的智能物体(Smart Object)、基于 Internet(Internet-based，Web-enabled)的无线传感器网络等应用的基本技术。由 27 个公司发起了针对智能对象联网的 IP 标准协作组织——IPSO(IP for Smart Object alliance)，目前已有 45 个成员，包括 Cisco、SAP、SUN、Bosch、Intel 等，该组织提出的 IPv6 协议栈 uIPv6 可以和主流厂商的协议栈互操作，其轻量级的代码只需要 11.5KB 的内存。

8.2.1 6LoWPAN 协议简介

IETF 于 2004 年成立 6LoWPAN(IPv6 over low-power Wireless Personal Area Network)工作组，致力于将 TCP/IPv6 协议栈构建于 IEEE 802.15.4 标准之上，并且通过路由协议构建起自组织方式的低功耗、低速率的 6LoWPAN 网络。第一个 6LoWPAN 规范 RFC4919 给出标准的基本目标和需求，然后在 RFC4944 中规范了 6LoWPAN 的格式和功能。通过部署和实现的经验，6LoWPAN 工作组进一步公布了包头压缩(Header Compression)、邻居发现(Neighbor Discovery)、用例(Use Case)及路由需求等文档。2008 年 IETF 成立了一个新的工作组：ROLL(Routing Over Low-power and Lossy Networks)，规范了低功耗有损网络(Low-power and Lossy Network，LLN)中路由的需求及解决方案。

在 6LoWPAN 提出后，很多组织、标准或联盟都提出了相应的兼容性方案。

在 2008 年，ISA 开始为无线工业自动化控制系统制定标准，称为 SP100.11a(也称为 ISA100)，该标准基于 6LoWPAN。

同样是 2008 年，IPSO(IP for Smart Objects)联盟成立，推动在智能物体上使用 IP 协议。

IP500 联盟主要致力于针对商业和企业建筑自动化和过程控制系统的开放无线 Mesh 网络，是一个在 IEEE 802.15.4 子(sub-GHz)无线电通信上建立 6LoWPAN 的联盟，sub-GHz ISM 频段是 433MHz、868MHz 和 915MHz，使用该频段的原因是当2.4GHz ISM 频段变得拥挤时，sub-GHz 比 2.4GHz 有更好的低频穿透能力，导致更大的传输距离。

开放地理空间论坛(Open Geospatial Consortium，OGC)规范了一个基于 IP 的地理空间和感知应用的解决方案。

2009 年，欧洲通信标准研究院(European Telecommunication Standards Institute，ETSI)成立了一个工作组，制定 M2M 标准，其中包括端到端的与 6LoWPAN 兼容的 IP 架构。如图 8.8 所示给出了 6LoWPAN 与相关标准和联盟的关系[6]。

物联网中特别是可通过 Internet 访问的传感器网络，其结点数目巨大，分布在户外并

且位置可能是动态变化的。IPv6 由于具有地址空间大、地址自动配置、邻居发现等特性，因此特别适合作为此类物联网的网络层。同时在技术上，IPv6 的巨大地址空间能够满足结点数量庞大的网络地址需求；IPv6 的一些新技术（如邻居发现、无状态的地址自动配置等技术）使自动构建网络时要相对容易一些。IPv6 与 IEEE 802.15.4 的 MAC 层的结合，可以轻松实现大规模传感器（智能物体）网络与 Internet 的互连，并能够远程访问这些传感器（智能物体）结点的数据。6LoWPAN 就是介于 IPv6 和 IEEE 802.15.4 之间的一个适配层，其协议栈如图 8.9 所示。

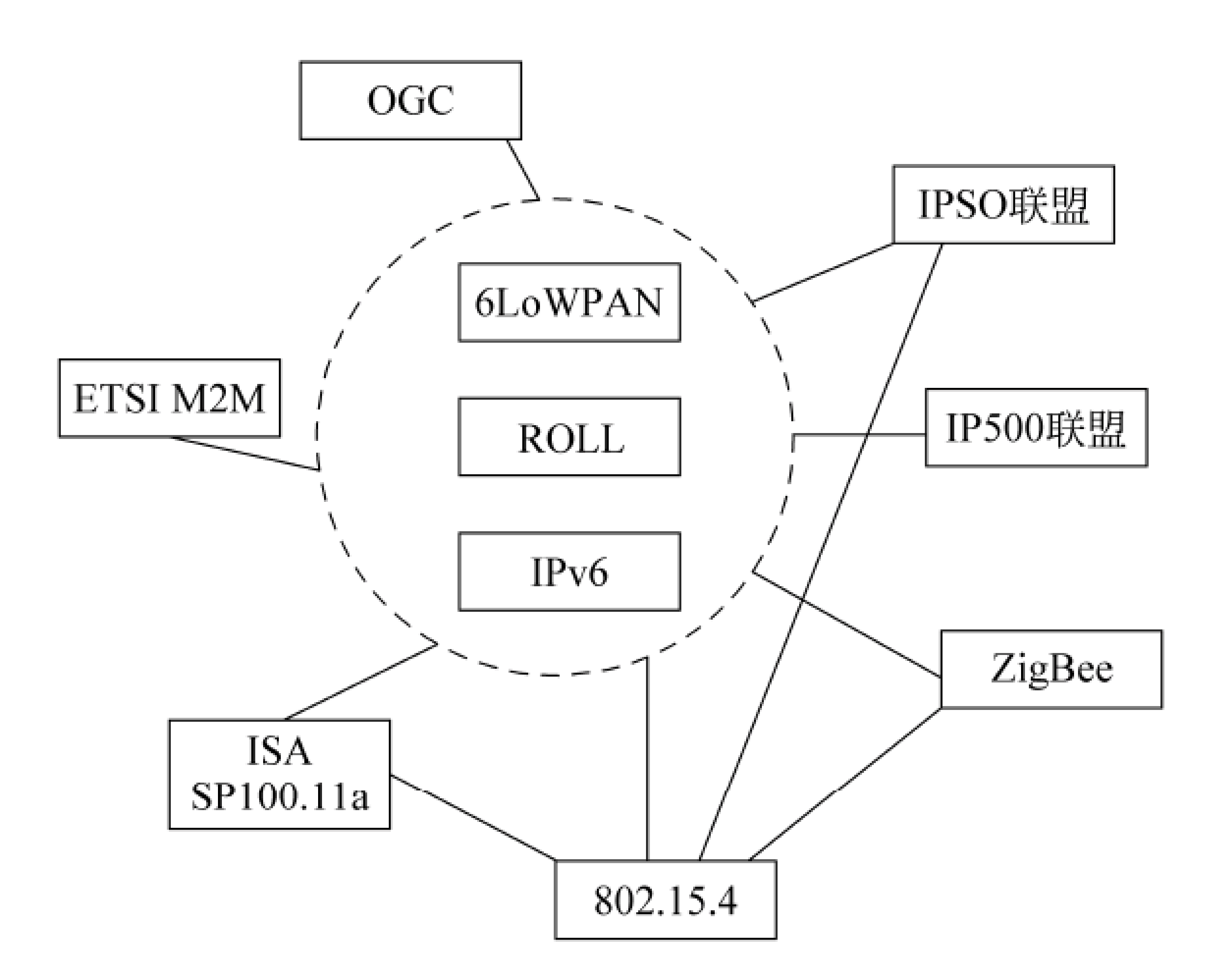


图 8.8　6LoWPAN 与相关标准和各工业联盟标准间的关系

应用层
传输层
精简IPv6协议层(MiniIPv6)
6LoWPAN适配层
IEEE 8 0 2.15.4 MAC层
IEEE 8 0 2.15.4 PHY层

图 8.9　6LoWPAN 协议栈参考模型

在构造物联网时，往往涉及传统 IP 网络和基于 IP 的 WPAN（无线个域网）的互连。具有 6LoWPAN 的协议栈和传统 IP 协议栈的比较如图 8.10 所示[6]。相应地，在传统 IP 网络和物联网之间的边界路由器上，需要实现两类数据包的处理和转发，于是其路由器协议栈如图 8.11 所示[6]。

IP协议栈				6LoWPAN协议栈		
HTTP	RTP		应用层	HTTP	RTP	
TCP	UDP	ICMP	传输层	TCP	UDP	ICMP
IP			网络层	IP		
以太网MAC			链路层	6LoWPAN 802.15.4MAC		
以太网PHY			物理层	802.15.4PHY		

图 8.10　IP 和 6LoWPAN 协议栈的比较

IPv6	
以太网MAC	6LoWPAN 802.15.4MAC
以太网PHY	802.15.4 PHY

图 8.11　支持 6LoWPAN 的 IPv6 边界路由器协议栈

8.2.2　6LoWPAN 要解决的问题

IETF 6LoWPAN 草案标准是专门为将 IP 扩展到低速率有损无线网络而设计的，其在整个 TCP/IP 协议栈中的位置如图 8.12 所示，是处于 IP 和 802.15.4 之间的一个适配

层。该适配层的功能包括包的分片/组装、试运行/启动(自动配置)、邻居发现的优化、Mesh 路由等功能。

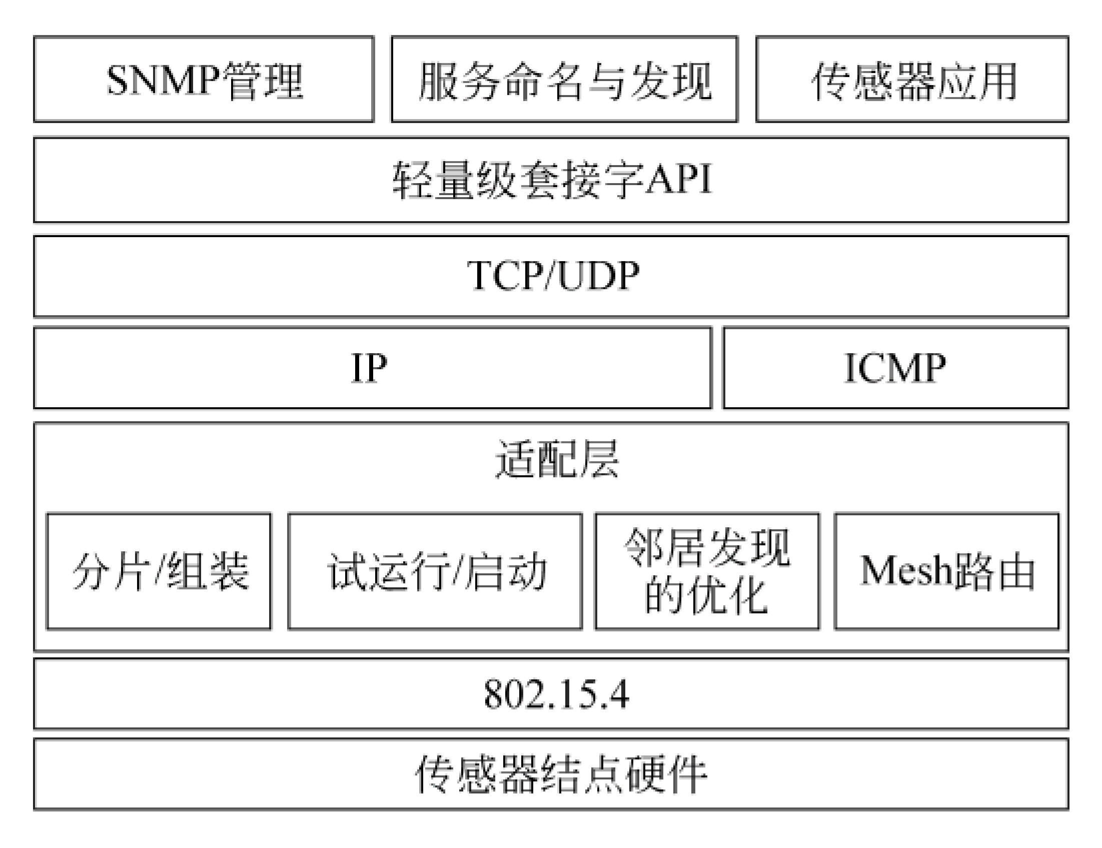


图 8.12　LoWPAN 结点协议架构

具体而言,6LoWPAN 需要解决的问题有如下。

(1) IP 连接问题：IPv6 巨大的地址空间和无状态地址自动配置技术使数量巨大的传感器结点(智能物体)可以方便地接入包括 Internet 在内的各种网络。但是,由于有报文长度和结点能量等方面的限制,标准的 IPv6 报文传输和地址前缀通告无法直接用于 IEEE 802.15.4 网络。

(2) 网络拓扑：IPv6 over IEEE 802.15.4 网络需要支持星形和 Mesh 拓扑。当使用 Mesh 拓扑时,报文可能需要在多跳网络中进行路由,类似于 Ad-Hoc 网络中的情形。但同样是由于报文长度和结点能量的限制,IEEE 802.15.4 上层路由协议应该更简单,管理的消耗也应该更少。此外,还需要考虑到结点计算和存储能力的限制。具体而言：路由协议对于数据报文的开销必须小,且与路由跳数无关;在路由过程中所需要的内存和计算必须小,以满足低开销低能量的目的;不可能使用大量的内存来维护路由表;在 6LoWPAN 支持的所有拓扑中,各类结点都能选择电池供电或者固定电源供电,因此需要考虑在休眠模式下路由协议的实现;能够通过网关或者其他方式同其他类型的网络(如以太网)进行无缝连接。

(3) 报文长度限制：IPv6 要求支持最小 1280 字节的 MTU,而 IEEE 802.15.4 最大支持 102 字节 MAC 帧长。一方面需要 IEEE 802.15.4 网络的应用尽量发送小的报文以避免分片,另一方面也需要结点在链路层提供对超过 102 字节的 IPv6 报文的分片和重组。

(4) 有限的配置和管理：在 IEEE 802.15.4 网络中,大量设备被期望能布置于各种环境中,而这些设备仅仅拥有有限的显示和输入功能。因此,IEEE 802.15.4 网络所使用的协议应该是只需要最少量配置并且易于初始化。有些部署地点是人无法到达的地方,因此需要结点有一定的自配置功能。另外 MAC 层以上运行的协议的配置也要尽量简单,并且需要网络拓扑有一定的自愈能力。

(5) 组播限制：IPv6 特别是其邻居发现协议的许多功能均依赖于 IP 组播。然而

IEEE 802.15.4 仅提供有限的广播支持,不论在星形还是 Mesh 拓扑中,这种广播均不能保证所有的结点都能收到封装在其中的 IPv6 组播报文。

(6) 安全问题:IEEE 802.15.4 提供基于 AES 的链路层安全支持,然而该标准并没有定义诸如初始化、密钥管理及上层安全性之类的细节。

对上述问题,6LoWPAN 工作组提出了一些解决方案,解决方案的性能评价指标主要是报文消耗、带宽消耗、处理需求及能量消耗,这四个方面也是影响 6LoWPAN 网络性能的主要因素。下面简介一下解决方案。

(1) 分片与重组:为了解决 IPv6 最小 MTU 为 1280 字节与 IEEE 802.15.4 Payload 长度仅有 81 字节冲突的问题,6LoWPAN 需要对 IPv6 报文进行链路层的分片和重组。

(2) 报头压缩:在使用 IEEE 802.15.4 安全机制时,IP 报文只有 81 字节的空间,而 IPv6 头部需要 40 字节,传输层的 UDP 和 TCP 头部分别为 8 和 20 字节,这就只留给了上层数据 33 或 21 字节。如果不对这些报头进行压缩的话,6LoWPAN 数据传输的效率将是非常低的。

(3) 组播支持:IEEE 802.15.4 并不支持组播也不提供可靠的广播,6LoWPAN 需要提供额外的机制以支持 IPv6 在这方面的需要。

(4) 网络拓扑管理:IEEE 802.15.4MAC 层协议仅提供基本的点对点的传输,无法很好地支持 IPv6。因此必须在 IP 层以下、MAC 层以上构建一定的网络拓扑,形成合适的拓扑结构,如星形、树形或者 Mesh。6LoWPAN 负责调用 MAC 层提供的原语,以形成正确的多跳拓扑。

(5) Mesh 路由:一个支持多跳的 Mesh 路由协议是必要的,但现有的一些无线网络路由协议(如 AODV 等)并不能很好地适应 LoWPAN 的特殊情况,这些路由协议大多是通过广播方式进行路由询问,对于能量供应相当有限的结点来讲很不现实。

(6) 安全性:6LoWPAN 需要考虑安全性。这一方面还有很多工作要开展。对应与上面的 6LoWPAN 引入的新处理,可能存在的安全问题包括:分片与重组攻击,报头压缩相关的攻击(如错误的压缩,拒绝服务攻击),轻量级组播安全,Mesh 路由安全等。

8.2.3 6LoWPAN 的安全性讨论

因为分片与重组的存在,报文中与分片/重组过程相关的参数有可能会被攻击者修改或重构,如数据长度(datagram_size)、数据标签(datagram_tag)、数据偏移(datagram_offset)等,从而引起意外重组、重组溢出、重组乱序等问题,进而使结点资源被消耗、停止工作、重启等,以这些现象为表现的攻击被称为 IP 包碎片攻击(IP packet fragmentation attack),进而可引发 DoS 攻击和重播攻击。所以,H. Kim 等人[11]提出了在 6LoWPAN 适配层增加时间戳(Timestamp)和随机序列(Nonce)选项来保证收到的数据包是最新的,从而防止数据包在传输过程被攻击者修改或重构,进而有效地防止 IP 包碎片攻击。

W. Jung 等人[12]提出并实现了一整套在 6LoWPAN 网络中实现 SSL(Secure Sockets Layer,安全套接层协议层)的方案,他们在密钥分发上对 ECC 和 RSA 做了比较,在密码算法上对 RC4、DES、3DES 做了比较,在消息认证上使用 MD5 和 SHA1 函数,最后发现 ECC-RC4-MD5 的组合消耗的资源最小,分别占用 64KB 的 flash 和 7KB 的

RAM，实现一次完整的 SSL 握手需要 2s。

RFC 工作文档给出了一些对 6LoWPAN 安全的分析。

关于 IEEE 802.15.4 的安全性。IEEE 802.15.4MAC 层提供了安全服务，由 MAC PIB(PAN Information Base)控制，MAC 子层在 PIB 中维护一个访问控制列表(ACL)。通过针对某个通信方设定一个 ACL 中的安全套件(Security Suite)，设备可以确定使用什么安全级别(即无安全、访问控制、数据加密、帧完整性等)与该通信方通信。

IEEE 802.15.4MAC 的一个关键功能就是提供了帧安全性。帧安全性其实是 MAC 层提供给上层的可选服务。取决于应用的需求，若应用并没有设定任何安全参数，则这一安全功能缺省是中止的。IEEE 802.15.4 定义了 4 种包类型：Beacon 包、数据包、确认包以及控制包。对于确认包没有安全机制。其他的包类型可以选择是否需要完整性保护或者保密性保护。由于 IEEE 802.15.4 的应用十分广泛，因此认证和密钥交换机制在标准中并没有定义，留给上层应用来定义。

关于 IP 的安全性，IPSec 可以保证 IP 包的完整性和保密性。IPSec 支持 AH 来认证 IP 头，以及 ESP 来认证和加密包负载。IPSec 的主要问题是处理能耗和密钥管理。目前并不清楚在 6LoWPAN 结点上实现 SADB、策略库、动态密钥管理协议是否是合适的。基于目前的硬件情况，6LoWPAN 结点上不适合实现所有的 IPSec 算法，即使是功能略强的 FFD 或者 RFD 结点上。另外，由于带宽也是 6LoWPAN 中一个非常紧缺的资源，IPSec 需要在每个包中额外传输包头(AH 或者 ESP)可能会带来沉重的负担。IPSec 需要两个通信方共享一个秘密密钥，这一密钥通常是通过 IKEv2 协议建立的，因此，这又增加了 IKEv2 协议的通信负担。由于邻居发现协议在 6LoWPAN 中使用，因此安全邻居发现协议(Secure Neighbor Discovery，SeND)应该被考虑。SeND 在 IP 网络中工作良好，但协议中使用的 CGA(Crypto-Generated Address)技术是基于 RSA 密码的，RSA 与椭圆曲线秘密(ECC)相比需要更大的包尺寸和处理时间。因此，一个合理的可能性就是在 SeND 协议中使用 ECC 来用于 6LoWPAN 网络。

关于密钥管理方面，指出由于结点资源受限，缺乏物理保护，无人值守操作，且于物理环境的密切交互，这些都使得在 6LoWPAN 中使用常用的密钥交换技术变得不太可行。常见的 3 种密钥管理技术，如基于可信第三方的密钥分配技术、密钥预分配技术、基于公钥密码的技术均面临一些困难。基于可信第三方的技术，如 Kerberos，具有单一失效点，这一方法不适合 6LoWPAN，因为不能保证和可信第三方的连接总是可用的，特别是在 LLN 网络中。基于密钥预分配的技术需要网络部署者事先知道结点的布局，结点之间的相邻关系，但是，由于结点部署的随机性，这种相邻关系可能无法事先获得。而且，若结点可能在网络部署是被入侵者攻击，动态在线(on-site)密钥管理技术比起密钥预分配要更加有利于处理网络的动态性。基于公钥密码的密钥分配技术，如数字证书，在 6LoWPAN 结点上可能计算能耗较高，如 DH 密钥协商、RSA 或者 ECC 等，但是有研究表明 ECC 可在传感器结点上实现。在密钥管理方面的建议包括：

(1) 敌对结点可能在结点布置阶段隐藏在其他结点之中，因此在启动阶段的安全密钥分配是一个研究问题。

(2) 结点在工作过程中被捕获，因此，密钥回收必须考虑。

（3）在睡眠模式中，给睡眠结点的密钥必须可以从唤醒模式的结点中推导出来。

（4）一旦密钥暴露了，必须诊断安全的破坏情况。

（5）密钥管理机制应该允许增加新的结点。

*8.2.4 RPL 和 CoAP 的安全性讨论

本章最后简要讨论一下在 6LoWPAN 层以上的路由协议 RPL 的安全性以及应用层 CoAP 协议的安全性。两者的安全性研究仍然在继续中。

1. RPL 的安全性

IPv6 的路由协议需要修改以适合 LLN，即 RPL（Routing Protocol for LLN）协议。该协议定义了能够在 LLN 环境中使用的点到点、点到多点、多点到点的路由协议。RPL 是一个高度模块化的协议，其路由协议的核心满足特定应用的路由需求的交集，而对于特定的需求，可以通过添加附加模块的方式满足。RPL 是一个距离向量协议，它创建一个 DODAG（Destination Oriented Directed Acyclic Graph），其中路径从网络中的每个结点到 DODAG 根（通常是汇点或者 LBR）。使用距离向量路由协议而不是链路状态协议的主要原因是低功耗有损网络中结点资源受限的性质。链路状态路由协议虽然更加强大，但是需要大量的资源，例如内存（链路状态数据库 LSDB）和用于同步 LSDB 的控制流量。

先简要介绍一下 RPL 中用到一些术语[15]。DAG（Directed Acyclic Graph）是有向非循环图。DAG Root 表示 DAG 根结点。所有的 DAGs 必须有至少一个 DAG 根，并且所有路径终止于一个根结点。DODAG（Destination Oriented DAG）是面向目的地的有向非循环图，以单独一个目的地为根的 DAG。DODAGRoot 是一个 DODAG 的 DAG 根结点。它可能会在 DODAG 内部担当一个边界路由器，尤其是可能在 DODAG 内部聚合路由，并重新分配 DODAG 路由到其他路由协议内。Rank 表示等级。一个结点的等级定义了该结点相对于其他结点关于一个 DODAG 根结点的唯一位置。OF（Objective Function）表示目标函数。定义了路由度量，最佳目的，以及相关函数如何被用来计算出 Rank 值。此外，OF 指出了在 DODAG 内如何选择父结点从而形成 DODAG。

下面简要介绍 RPL 的基本过程。首先，网络管理员把一个或多个结点配置为 DODAG 根。RPL 使用了新定义的 ICMPv6 消息的结点发现机制来创建 DODAG。RPL 定义了两种新的 ICMPv6 消息，叫做 DODAG 信息对象（DIO）消息和目的地通过消息对象（DAO）消息。DIO 消息是由结点发送的，用于通告有关 DODAG 的消息，例如 DODAGID、OF、DODAG 级别、DODAG 序列号以及其他 DODAG 参数。当一个结点发现多个 DODAG 邻居，它会使用多种规则来决定是否加入 DODAG。一旦一个结点加入到一个 DODAG 中，它就会拥有到 DODAG 根的路由，用来支持从叶结点到 DODAG 根的 MP2P 流量。

RPL 支持 3 种安全模式，即不安全模式、预置安装模式和授权模式，但在低功耗有损网络中实施这些模式的复杂性是需要首先关注的问题。

2. CoAP 的安全性简介

CoAP（Constrained Application Protocol）协议是用于 M2M 应用的轻量级应用层协议，可以作为智能物体网络的应用层协议。文献[18]给出了 CoAP 安全架构方面的讨论，

这里简要介绍一下。

该文中的安全架构同时也是一个部署模型，这一架构的根本点是自生成安全标识(self-generated secure identity)，这与 CGA(Cryptographically Generated Addresses)类似。即令 I=h(P‖O)，其中 I 是设备的安全标识，h 是 Hash 函数，P 是设备生成的公钥，O 是可选的其他信息。安全标识可用于安全凭证、共享的秘密、安全策略信息。安全标识可用于识别认证的设备。有多种方式可完成在设备部署阶段收集标识信息。例如，可以将标识的最后几位数字打印在设备上。或者设备的包装上包含整个标识，这一标识可以通过条形码读取。无论收集的方式如何，这一模型方便了安全设置。每个设备在一个随机的安全的密钥生产过程中生成自己的标识，这一标识是自安全的，也就是说如果知道通信对方的标识，来自对方的消息可以被对方的私钥所签署，非常容易验证来自所期望的对方的消息的真实性。没有必要分别配置标识和该标识的证书，也没有必要配置组密钥或者共享密钥。

研究与思考

[1] 论证 IPSec 和 SSL 机制能否在资源受限结点(如传感器结点)上实现。可参考进一步阅读建议的文献[1,2,3]，作为一个实践项目来完成。

[2] 如何在 SeND 协议中实现 ECC 算法来支持 6LoWPAN 网络。

[3] 如何给出 RPL 的安全性分析。

[4] 如何给出 CoAP 的安全性分析。

[5] 我国提出的 IPv9 协议最近获得了美国专利，有人认为它使得“我国是目前世界上唯一能实现域名、IP 地址和 MAC 地址统一成十进制文本表示方法的国家。同时，也成为继美国之后，第二个在世界上拥有根域名解析服务器和 IP 地址硬连接服务器的国家和世界上第二个拥有自主的域名、IP 地址和 MAC 地址资源的国家及可独立进行域名解析和 IP 地址硬连接，并可独立自主的分配域名、IP 地址和 MAC 地址的国家。”(摘自百度百科 IPv9。)但有的人说它不过是个愚人节玩笑。IPv9 充满了争议，你认为它是否合理，对我国的网络安全是否有意义。给出自己的调研、分析和结论。

进一步阅读建议

这里给出部分图 8.8 涉及的标准或者联盟组织。

[1] Jun-Cheol Park, Ah-Hyun Jun, A lightweight IPsec adaptation for small devices in IP-based mobile networks[C], In Proc. of The 8th International Conference Advanced Communication Technology (ICACT06), 298-302, 20-22 Feb. 2006.

[2] S. Raza, S. Duquennoy, T. Chung, D. Yazar, T. Voigt, U. Roedig, Securing communication in 6LoWPAN with compressed IPsec[C], In Proc. of 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), 1-8, 27-29 June 2011.

[3] Wooyoung Jung, Sungmin Hong, Minkeun Ha, Young-Joo Kim, Daeyoung Kim, SSL-Based Lightweight Security of IP-Based Wireless Sensor Networks[C], In Proc. of International

Conference on Advanced Information Networking and Applications Workshops(WAINA09), 1112-1117, 26-29 May 2009.
[4] IETF 6LoWPAN: Header Compression[S], http://tools.ietf.org/html/draft-ietf-6lowpan-hc-08.
[5] IETF 6LoWPAN: Neighbor Discovery[S], http://tools.ietf.org/html/draft-ietf-6lowpan-nd-11.
[6] IETF Routing over Low-power and Lossy Networks (ROLL)[S], http://tools.ietf.org/ html/draft-ietf-roll-rpl-10.
[7] Constrained Application Protocol (CoAP)[S], http://tools.ietf.org/html/draft-ietf-core-coap-01.
[8] IETF Smart Power Interest Group[S], https://www.ietf.org/mailman/listinfo/smartpower-interest.
[9] IP500[OL], http://www.ip500.de/.
[10] ETSI M2M Standardization[S], http://www.etsi.org/.

本章参考文献

[1] IPSO 联盟[OL], http://ipso-alliance.org/.
[2] 刘宗伟. IPv6 安全技术研究[D]. 吉林大学硕士学位论文, 2008.
[3] 王杰. 计算机网络安全的理论与实践(第 2 版)[M]. 北京: 高等教育出版社, 2011.
[4] Behrouz A. Forouzan, 马振晗, 贾军保译. 密码学与网络安全[M]. 北京: 清华大学出版社, 2009.
[5] 冯登国. SSL/TLS 安全协议就此寿终正寝了吗? [J]. 信息安全与通信保密, 2010, (04).
[6] Z. Shelby, C. Bormann, 6LoWPAN: The Wireless Embedded Internet [M], John Wiley & Sons, Nov. 2009.
[7] P. Thubert, IETF ROLL WG, RPL: IPv6 Routing Protocol for Low power and Lossy Networks [S], May 28, 2010, http://tools.ietf.org/html/draft-ietf-roll-rpl-08.
[8] N Kushalnagar, G Montenegro, C Schumacher, IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals [S], IETF RFC4919, 2007.
[9] G. Montenegro, N. Kushalnagar, J. Hui, D. Culler, Transmission of IPv6 Packets over IEEE 802.15.4 Networks [S], IETF RFC 4944, September 2007, http://www.ietf.org/rfc/rfc4944.txt.
[10] Routing Over Low power and Lossy networks (ROLL) Working Group [OL], https://datatracker.ietf.org/wg/roll/.
[11] H. G Kim. Protection against Packet Fragmentation Attacks at 6LoWPAN Adaptation Layer[C]. In Proc. of International Conference on Convergence and Hybrid Information Technology, 2008.
[12] W. Jung, S. Hong, M. Ha, SSL-based lightweight security of IP-based wireless sensor networks [J]. In Proc. of 2009 International Conference on Advanced Information Networking and Applications Workshops, 2009.
[13] S. Park et al., IPv6 over Low Power WPAN Security Analysis [S], draft-daniel-6lowpan-security-analysis-04.txt, http://tools.ietf.org/html/draft-daniel-6lowpan-security-analysis-04.
[14] 刘外喜, 唐冬, 胡晓, 郑晖. 6LoWPAN 网络安全问题的分析[J]. 电信科学, 2010(4).
[15] 宋菲, 侯乐青. 浅析智能物件网络中的 RPL 路由技术[J]. 电信网技术, 2011/09.
[16] Park, S., Kim, K., Haddad, W., Chakrabarti, S., and J. Laganier, IPv6 over Low Power

WPAN Security Analysis [S], draft-daniel-6lowpan-security-analysis-05 (work in progress), March 2011. http://tools.ietf.org/html/draft-daniel-6lowpan-security-analysis-05.

[17] C. Bormann, A. P. Castellani, Z. Shelby, CoAP: An Application Protocol for Billions of Tiny Internet Nodes[J], IEEE Internet Computing, vol. 16, no. 2, pp. 62-67, March-April 2012.

[18] Shelby, Z., Hartke, K., Bormann, C., and B. Frank, Constrained Application Protocol (CoAP) [S], draft-ietf-core-coap-06 (work in progress), May 2011.

[19] J. Arkko, A. Keranen, CoAP Security Architecture[OL], July 26, 2011, http://tools.ietf.org/html/draft-arkko-core-security-arch-00.

第 9 章　物联网服务端安全——云计算安全

前面章节讨论了物联网网络层的接入网安全和核心网安全，本章介绍服务端的安全问题。服务端安全涉及的范围可以更广，如服务器端的访问控制技术、数据库安全相关技术、P2P 安全技术等，但本章选择介绍的是一种新兴的热门服务端模式，即云计算。这是因为云计算与物联网相结合，可以发挥"物"端和"云"端的各自特点和优势，"物"端的轻便性往往制约了端设备的存储和计算能力，"云"端可以弥补这一不足，提供柔性按需存储和计算能力。特别是当物联网的规模足够大时，就需要和云计算结合起来，例如，在大型的行业应用中需要大量的后端数据支持和管理；M2M 应用中接入网络的终端数量规模巨大，因此都需要云计算中心提供强大的后端存储和计算支持。

由于云计算的应用越来越受到重视，如阿里云、百度云、华为云、iCloud、Azure 等商用云计算服务的推出，可见非计算机企业已经开始涉足云计算，尤其是传统的通讯企业华为，也开始涉足这一领域。2012 年世界移动通信大会上，华为进行了世界首个移动宽带加速云解决方案的现场演示，该方案通过端（终端）、管（网络）、云（服务）各个数据传送环节的创新优化，令数据终端的访问速度大幅提升 30%～80%。另外，微软 CEO 鲍尔默称 Windows 8 将借云计算重塑软件帝国，Windows 8 推云服务 SkyDrive 将对抗甚至击败 iCloud。因此，随着云计算应用的普及，云计算的安全性也受到越来越多的关注。

9.1　云计算及其安全问题

9.1.1　云计算简介

云计算（Cloud Computing）是网格计算（Grid Computing）、分布式计算（Distributed Computing）、并行计算（Parallel Computing）、效用计算（Utility Computing）、网络存储（Network Storage Technologies）、虚拟化（Virtualization）、负载均衡（Load Balance）等传统计算机和网络技术发展融合的产物。

美国 NIST 对云计算的定义是[1]：云计算是一种按使用计费（Pay-per-use）模型，提供对可安装且可靠的计算资源共享池进行方便按需的网络访问服务，如网络、服务器、存储、应用、服务等。这些资源能够快速地供给和发布，仅需要最少的用户管理和服务供应商间交互。云是一个容易使用且可访问的虚拟资源池，比如硬件、开发平台和服务。这些资源可进行动态重安装以适应变化的负载，允许资源的优化利用，这个资源池通常付费使用，其质量通常是通过基础设施供应商与客户间的服务层共识（Service Level Agreement）来保障的。因此，云计算的 5 个关键特征是：按需自服务（On-Demand Self-service），普适网络访问（Uniquitous Network Access），资源池（Resource Pooling），快速

弹性(Rapid Elasticity),按使用计费。

云计算的体系结构包括 3 个部分:应用层 SaaS(Software as a Service)、平台层 PaaS (Platform as a Service)、基础设施层 IaaS(Infrastructure as a Service),如图 9.1 所示,这里特意添加了中国云计算领域的公司。

SaaS(Software as a Service): SalesForce、Microsoft Live Office、道里可信云安全
PaaS(Platform as a Service): Google App Engine、Microsoft Azure、Force.com、友友系统
IaaS(Infrastructure as a Service): Amazon EC2、世纪互联CloudEx、天云科技
Data Centers (Data Farm, Servers)、Networks、Hardwares

图 9.1　云计算的体系结构

(1) 基础设施层提供对计算、存储、带宽的管理,是虚拟化技术、负载均衡、文件系统管理、高可靠性存储等云计算机关键技术的集中体现层,是平台服务和应用服务的基础。

(2) 平台层为应用层的开发提供接口 API 调用和软件运行环境。应用层开发调用 API,不需要考虑具体负载均衡、文件系统、储存系统管理等实现细节。

(3) 应用层服务提供具体应用,是一种通过 Internet 提供软件的模式,如 SalesForce 的客户关系管理(CRM)、Google Apps 等。

根据云基础设施的规模和服务对象对云进行分类。NIST 认为云有两种类型:内部云和外部云,以及 4 种部署模型:私有云、社区云、公共云、异种云。维基百科认为可分为私有云、公共云和异种云。

(1) 私有云或企业云(Private Cloud 或 Enterprise Cloud):主要是大型企业内部拥有的云计算数据中心,如银行、电信等行业用户以及关注数据安全的用户。大型企业的 IT 部门无需将业务完全转给公共云供应商,他们会保留原有系统,但新增系统将选用基于云计算的架构。

(2) 公共云(Public Cloud):云计算基础设施供应商拥有的大量数据中心,为中小企业提供云平台和云应用,即通常所指的公共云计算供应商。为社区服务的公共云可视为社区云。公共云中专门为某个企业服务的云可视为托管云。

(3) 异种云或联邦云(Hybrid Cloud 或 Federal Cloud):提供云间的互操作接口,各种云的集合体。如 VMWare VCloud、OpenNebula。

云计算具有如下的优势:

(1) 协同工作方便。通过云计算的应用层,即 SaaS 层,可方便地在多个用户间协同工作。比如使用 Google Docs 协同编辑文档,使用 SalesForce 管理商业活动等。

(2) 无时无处不在的享受服务。云计算将融合各种现有计算资源,如对等计算(P2P)、网格计算、Web 服务等。通过各种有线或无线接入网络,如 ADSL、Wi-Fi、WiMAX、3G、卫星网络等,为各种客户端,如智能手机、PDA、车辆网络结点、体域网、传感器网络、可穿戴计算(Wearable Computing)、智能信息家电的嵌入式系统等,提供真正意

义的普适(Uniquitous)计算和服务。

(3) 系统可扩展性,伸缩性较高。当需求增大时,无需购买新的设备并升级硬件资源,系统自动升级或暂时支付服务租用的租金。这是因为云计算使用虚拟的计算资源,当资源不够用时,可以通过增加虚拟资源的方法无缝升级系统(如使用 PowerVM、Xen 增加一个计算实例)。这特别适合应对可能出现短暂高峰资源请求的情况。

(4) 系统可用性,可靠性高。由于资源是高度分布和虚拟化的系统,在构建时通过专业手段(如数据备份、系统冗余)保证高可靠性和高可用性,可以保障 24×7 小时的不间断服务。

(5) 中小企业节约 IT 成本。使用云计算的好处在于初期只需要很小投资即可使用服务。由于云计算使用根据时间和流量的付费,硬件和软件的费用全免,所以初期投资与传统方式自建(购)系统相比要小很多。同时,也节约了系统维护的人力成本。云计算可帮助企业节约大约 80%的使用面积,60%的电源和制冷消耗,达到 3 倍的设施利用率。

9.1.2 云计算的安全问题

云计算的安全问题应包括 3 个主要方面:信任(Trust)问题、网络与系统安全问题(Security)、隐私保护(Privacy)问题。信任问题包括云服务的信任评价、信任管理等问题。云计算的网络安全问题包括云计算数据传输的通信安全问题;系统安全问题,如云计算平台的可靠性问题;数据存储安全问题等。其中数据存储安全是云计算应用服务能否被用户所接受和信赖的前提,也是迫切需要研究解决的问题。包括数据是否需要加密存储,如何加密,是在客户端还是服务器端加密,如何在不信任存储服务器的情况下保证数据存储的保密性和完整性,如何检查存储在云存储空间的数据完整性。另外,数据的隐私保护也十分关键,关系到客户是否愿意采用这一计算模式,包括用户的行为、兴趣取向等无法被推测。

云安全联盟 CSA(Cloud Security Alliance)[2]提出一个根据云计算的架构建立的云安全参考模型,如图 9.2 所示。由于参与者大多来自企业界,其视角比较侧重应用。该模型根据云计算的体系结构,从产品开发的视角,涵盖了网络安全和系统安全等,建立相应的安全保护机制,将现有网络安全机制根据云计算的新体系结构做了相应的调整。

我国著名信息安全专家冯登国教授在文献[3]中给出了云计算安全的详细综述。认为云计算安全具有 3 个挑战:建立以数据安全和隐私保护为主要目标的云安全技术框架;建立以安全目标验证、安全服务等级测评为核心的云计算安全标准及其测评体系;建立可控的云计算安全监管体系。

云用户的安全目标主要有两个:数据安全与隐私保护服务,防止云服务商恶意泄漏或出卖用户隐私信息,或者对用户数据进行搜集和分析,挖掘用户隐私数据;安全管理,即在不泄漏其他用户隐私且不涉及云服务商商业机密的前提下,允许用户获取所需安全配置信息以及运行状态信息,并在某种程度上允许用户部署实施专用安全管理软件。

云安全服务可以分为可信云基础设施服务、云安全基础服务以及云安全应用服务 3 类。

可信云基础设施服务为上层云应用提供安全的数据存储、计算等信息资源服务。包

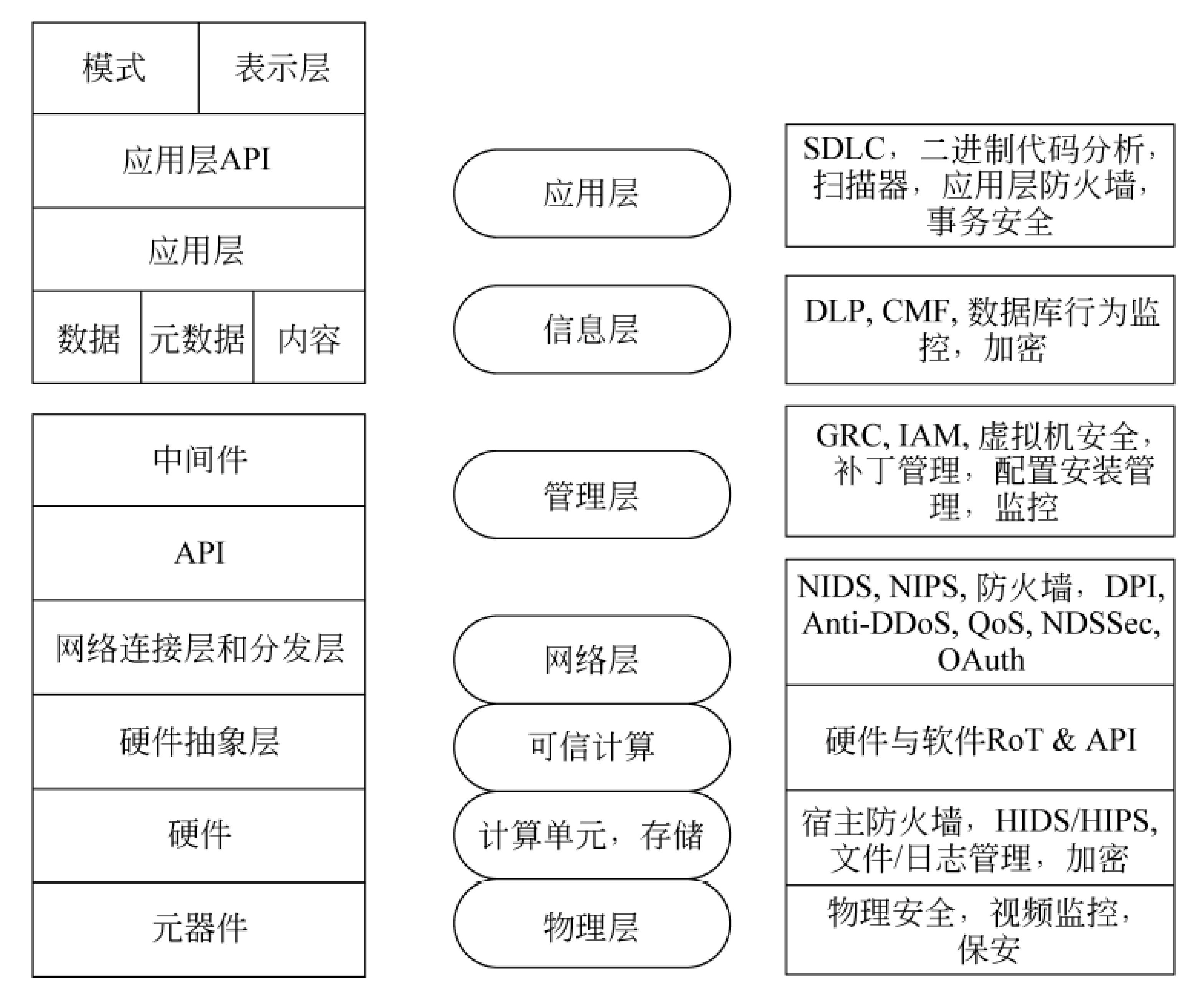


图 9.2　云安全参考模型

括两个方面：一个是云平台应分析传统计算平台面临的安全问题，采取全面严密的安全措施。例如在物理层考虑厂房安全，在存储层考虑完整性和文件/日志管理、数据加密、备份、灾难恢复等，在网络层考虑拒绝服务攻击、DNS 安全、网络可达性、数据传输机密性等，系统层则应涵盖虚拟机安全、补丁管理、系统用户身份管理等安全问题，数据层包括数据库安全、数据的隐私性与访问控制、数据备份与清洁等，而应用层应考虑程序完整性检验与漏洞管理等。另一方面，云平台应向用户证明自己具备某种程度的数据隐私保护能力。例如存储服务中证明用户数据以加密形式保存。计算服务中证明用户代码运行在受保护的内存中等。

云安全基础服务属于云基础软件服务层，为各类云应用提供共性信息安全服务，是支撑云应用满足用户安全目标的重要手段。其中比较典型的几类云安全服务包括云用户身份管理服务、云访问控制服务、云审计服务、云密码服务。

云安全应用服务与用户的需求紧密结合，种类繁多。典型的例子如 DDoS 攻击防护云服务、Botnet 检测与监控云服务、云网页过滤与杀毒应用、内容安全云服务、安全事件监控与预警云服务、云垃圾邮件过滤及防治等。

总的来说，云计算的安全研究问题应该主要是那些跟云计算的特征密切相关的新产生的安全问题，包括 10 个方面的具体问题：

(1) 数据存储安全问题：数据的完整性和保密性。由于数据存储在“云”端，且通常“云”端是不被信任的，因此需要保证托管的数据的完整性和保密性。

(2) 访问控制：服务访问控制策略的描述、访问控制的授权机制。

(3) 可信虚拟计算问题：包括安全的虚拟化计算、安全的虚拟进程移植、进程间安全

隔离等。

（4）信任管理：服务提供者之间的信任建立与管理、服务者与用户间的信任建立与管理。

（5）存储可靠性问题：将数据托管或者外包到“云”端存储，因此要注意数据分布式虚拟存储的健壮性和可靠性、存储服务的可用性、灾难恢复。

（6）鉴别与认证：用户标识管理，用户身份的认证。

（7）密钥管理：数据加密的密钥管理。

（8）加密解密服务：在何处进行数据的加密和解密，能否通过服务提供安全。

（9）云服务的安全：尤其 Web 服务的安全评估、安全扫描和检测。

（10）其他的问题：云计算的电子取证、云计算风险评估和管理、云供应商的规则遵守（Compliance）审计等。

数据存储安全和计算虚拟化安全是云计算两个急需解决的安全问题，后面将重点讨论这两个问题及其关键技术。

9.2 云计算的存储安全

云计算的数据存储安全问题通常包括如下 3 个方面：

（1）数据的访问控制问题。如何在数据加密的状态下进行访问控制，如何面对大规模海量数据进行高效的访问控制，如何在不信任云端服务的情况下确保访问控制。

（2）数据保密性问题，如何保障用户的数据不被泄漏或被云计算供应商窥探。如果这个问题没有解决，云计算便不可能存储关键或者敏感数据。用户的数据如何加密，在云端还是客户端加密，还是通过可信第三方加密。是否可能提供云加密服务，如何定义和设计这种服务。

（3）数据完整性问题，包括存储的数据不被非法修改，数据不丢失，以及存储的可靠性问题。数据的完整性如何验证，特别是客户端没有数据、同时又不信任云端存储的数据的情况下，验证数据的完整性。如果保存的数据频繁地更新，数据的完整性验证便更加困难。

云计算的实际安全需求，为密码学的发展提供了驱动力，最近几年的密码学学术会议上，常可以看到可应用于云计算特别是云存储安全而提出的密码学方法。有些方法甚至是密码学原语级的构造。

9.2.1 云存储的访问控制——基于属性的加密和代理重加密

针对云计算存储数据的访问控制，一个典型的需求就是由于数据是加密的，访问控制针对的对象是加密后的数据。通常的访问控制模型是基于角色的访问控制，即按照特定的访问策略建立若干角色，通过检查访问者的角色，控制访问者对数据的访问。但是该模型通常用于没有加密的数据，或者访问控制的控制端是可信的。若对于加密的数据采取这种访问控制，则需要使用将来欲访问该数据的用户的公钥去加密数据加密密钥（即异种

加密方式),这样的访问控制涉及大量的在数据上传客户端的加密运算、访问控制策略简单且不够安全。

一个非常新颖的想法就是,是否可能在加密的时候,已经将访问控制策略融入密文中,满足访问控制策略的用户在将来就可以正确地解密出密文。实现这一想法可依赖密文策略的基于属性的加密。

Sahai 与 Waters 在 2005 年提出基于属性的加密体制,发展了传统的基于身份密码体制中关于身份的概念,将身份看做是一系列属性的集合,提出了基于模糊身份的加密,将生物学特性直接作为身份信息应用于基于身份的加密方案中。基于模糊身份的目的是因为有些情况下需要只要大致具有该身份(属性)的人便可以解密数据,如医疗急救情形下的病患。2006 年 Goyal 等人在基于模糊身份加密方案的基础上提出了密钥策略基于属性的加密方案(Key Policy Attribute-based Encryption,KP-ABE)。2007 年,Bethencourt 等人提出了密文策略的基于属性的加密方案(Cipertext-policy ABE,CP-ABE),将用户的身份表示为一个属性集合,而加密数据则与访问控制结构(访问控制策略)相关联,一个用户能否解密密文,取决于密文所关联的属性集合与用户身份对应的访问控制结构是否匹配。

CP-ABE 的模型包括 4 个基本算法:Setup(.)、Encrypt(.)、KeyGen(.)及 Decrypt(.),简单描述如下:

(1) 参数生成算法 Setup:生成公开参数 PK 以及主密钥 MK。

(2) 加密算法 CT=Encrypt(PK,M,A):输入参数包括 PK、被加密的数据 M 以及访问控制策略 A。输出为密文 CT,CT 只能由那些具有满足访问控制策略 A 的用户才能解密。可见,在加密时已经将访问控制策略“嵌入”到密文中。

(3) 密钥生成算法 SK=KeyGen(MK,S):算法输入主密钥 MK 以及描述密钥的属性集 S,输出解密密钥 SK。可见,解密密钥和其是否满足访问控制策略的属性相关。

(4) 解密算法 M=Decrypt(PK,CT,SK):输入公共参数 PK,密文 CT 以及密钥 SK,当且仅当 S 满足访问控制策略 A,由属性集 S 产生的私钥 SK 才能解密 CT,此时,算法返回明文消息 M。

云存储的访问控制中也可以利用代理重加密机制。代理重加密就是指通过半可信的代理服务器(即相信它会按照规定的操作流程完成既定的工作,但是又不能让它知道明文),将本来是 A 的公钥加密的密文,重新加密成用 B 的公钥加密的密文。当然,A 事先在该代理服务器上设置 A 到 B 的重加密密钥 $K_{A\to B}$,代理服务器正是利用该重加密密钥将以 A 公钥加密的密文重加密成以 B 公钥加密的密文。在这一过程中,明文和 A 与 B 的私钥都不会暴露给代理服务器。

利用代理重加密机制 A 可以通过代理服务器重新分发密文给需要共享的用户,且保持明文保密性。它同时也可以利用云服务器计算能力强的特点,将加密的工作转移到服务器端完成。

9.2.2 云存储的数据保密性——同态加密 HE

为了保证数据的保密性,云存储端通常存储的是加密过的数据,为了对这些数据进行

操作，由于对云存储服务器不是完全信任，因而常规的办法是将这些数据发回到客户端，由客户端解密，然后进行相应的计算，完成后再加密上传到云存储端。这带来了很大的通信开销。是否有可能在云计算的存储端就能进行数据的计算，且不需要解密即针对密文进行呢，这时候就可以借助同态加密的思想。

同态加密(Homomorphic Encryption)是指对两个密文的操作，解密后得到的明文，等同与两个原始明文完成相同的操作。现有的多数同态加密算法要么只对加法同态(例如 Paillier 算法)，要么是只对乘法同态(例如 RSA 算法)，或者同时对加法和简单的标量乘法同态(例如 Iterated Hill Cipher、IHC 算法和 Modified Rivest's Scheme、MRS 算法)。虽有几种算法能够同时对加法和乘法同态(例如 Rivest 加密方案)，但存在严重的安全问题[11]。

2009 年，IBM 研究员 Craig Gentry 在计算机理论的著名会议 STOC 上发表论文，提出一种基于理想格(Ideal Lattice)的全同态加密算法，成为一种能够实现全同态加密所有属性的解决方案。虽然该方案由于同步工作效率有待改进而未能投入实际应用，但是它已经实现了全同态加密领域的重大突破。全同态加密能够在没有解密密钥的条件下，对加密数据进行任意复杂的操作，以实现相应的明文操作。本小节对同态加密进行简介。

设 x 和 y 是明文空间 M 中的元素，o 是 M 上的运算，$E_K()$是 M 上密钥空间为 K 的加密算法，称加密算法 $E_K()$对运算 o 是同态的，如果存在一个有效的算法 A，使得

$$A(E_K(x),E_K(y))=E_K(xoy)$$

不同运算的同态加密的简单形式化描述如下：

加法同态：给定 $E_K(x)$和 $E_K(y)$，存在一个计算上有效的算法 ADD 使得

$$E_K(x+y)=ADD(E_K(x),E_K(y))$$

即 $E_K(x+y)$可通过 $E_K(x)$和 $E_K(y)$轻易地计算出来，而不需要知道 x 和 y。

数量乘法同态：给定 $E_K(x)$和常数 t，存在一个计算上有效的算法 SMUL，使得

$$E_K(tx)=SMUL(E_K(x),t)$$

即 $E_K(tx)$可通过 $E_K(x)$和 t 轻易地计算出来，而不需要知道 x。

乘法同态：给定 $E_K(x)$和 $E_K(y)$，存在一个计算上有效的算法 MUL，使得

$$E_K(xy)=MUL(E_K(x),E_K(y))$$

即 $E_K(xy)$可通过 $E_K(x)$和 $E_K(y)$轻易地计算出来，而不需要知道 x 和 y。

其实，同态加密还可以运用到隐私保护的数据聚集(aggregation)上，例如智能电网中对智能电表的数据收集、无线传感器网络中感知数据的聚集等。

通常公钥加密方案 ε 由三个算法构成，$KeyGen_\varepsilon$、$Encrypt_\varepsilon$、$Decrypt_\varepsilon$，这三个算法必须都是可以高效计算的。也就是说，这三个算法的运行时间是安全参数 λ 的多项式时间 poly(λ)，其中，λ 通常刻画为密钥的比特长度。这 3 个算法描述如下：

$KeyGen_\varepsilon(\lambda)$：使用 λ 生成两个密钥，公钥 pk 与私钥 sk，即$(pk,sk)\leftarrow KeyGen_\varepsilon(\lambda)$。

$Encrypt_\varepsilon(pk,m)$：给定公钥 pk，将消息 m 映射为对应的密文 c，即 $c\leftarrow Encrypt_\varepsilon(pk,m)$。

$Decrypt_\varepsilon(sk,c)$：给定私钥 sk，将密文 c 映射为原始的消息 m，即 $m=Decrypt_\varepsilon(sk,c)$。

除了上述三个基本算法之外，同态公钥加密方案还包括第四个算法——$Evaluate_\varepsilon$，$Evaluate_\varepsilon$ 算法与一组功能函数 F_ε 相关联。对于 F_ε 中的每一个函数 f，以及任意密文 c_1，c_2，…，c_t，其中 $c_i \leftarrow Encrypt_\varepsilon(pk, m_i)$，算法 $Evaluate_\varepsilon(pk, f, c_1, \cdots c_t)$ 输出 $f(m_1, \cdots, m_t)$ 在公钥 pk 作用下对应的密文 c。也就是说，有 $Decrypt_\varepsilon(sk, c) = f(m_1, \cdots, m_t)$ 成立。显然，同态加密方案的一个基本条件就是 $Evaluate_\varepsilon$ 输出的密文能够被正确解密。

严格地说，方案需要具有正确性。即对于任意给定安全参数 λ，$KeyGen_\varepsilon(\lambda)$ 输出的任意密钥对(pk，sk)，对任意 $f \in F_\varepsilon$，给定任意明文 $m_1, m_2, \cdots, m_t$ 与对应的密文 $\vec{c} = (c_1, c_2, \cdots, c_t)$，其中 $c_i \leftarrow Encrypt_\varepsilon(pk, m_i)$，若 $c \leftarrow Evaluate_\varepsilon(pk, f, \vec{c})$，则必有 $f(m_1, \cdots, m_t) = Decrypt_\varepsilon(sk, c)$ 成立。

如果对于一类特定函数 F_ε 中的每一个函数 f，$Evaluate_\varepsilon$ 输出的密文都满足正确性要求，则加密方案 $\varepsilon = (KeyGen_\varepsilon, Encrypt_\varepsilon, Decrypt_\varepsilon, Evaluate_\varepsilon)$ 称为同态加密方案。

如果对于所有的布尔函数 f，$Evaluate_\varepsilon$ 输出的密文都满足正确性要求，则加密方案 $\varepsilon = (KeyGen_\varepsilon, Encrypt_\varepsilon, Decrypt_\varepsilon, Evaluate_\varepsilon)$ 称为全同态加密方案(Full Homomorphic Encryption，FHE)。

对于任意一个 FHE 方案而言，除了必需满足正确性条件外，还必须同时满足另外两个基本条件：保密性和紧凑性。

保密性。对于任意给定安全参数 λ，$KeyGen_\varepsilon(\lambda)$ 输出的任意密钥对(pk，sk)，对于任意 $f \in F_\varepsilon$，给定任意明文 $m_1, m_2, \cdots, m_t$ 与对应的密文 $\vec{c} = (c_1, c_2, \cdots, c_t)$，其中 $c_i \leftarrow Encrypt_\varepsilon(pk, m_i)$，给定 $c \leftarrow Evaluate_\varepsilon(pk, f, \vec{c})$，则除了密文输出 c 之外，$Evaluate_\varepsilon(pk, f, \vec{c})$ 不会泄露关于计算过程 f 的任何信息。也就是说，任何人无法获得关于函数 f 的任何信息，即使拥有解密密钥 sk。

紧凑性。对于任意给定安全参数 λ，$KeyGen_\varepsilon(\lambda)$ 输出的任意密钥对(pk，sk)，对于任意 $f \in F_\varepsilon$，给定任意明文 $m_1, m_2, \cdots, m_t$ 与对应的密文 $\vec{c} = (c_1, c_2, \cdots, c_t)$，其中 $c_i \leftarrow Encrypt_\varepsilon(pk, m_i)$，$c \leftarrow Evaluate_\varepsilon(pk, f, \vec{c})$，则 c 的长度不能超过某个有界多项式 b(λ)。换言之，即 c 的长度完全独立于 f。通俗地讲，c 看起来与任意一个普通密文 c_i 类似。

*9.2.3　云存储的数据完整性检验 POR 和 PDP

云存储服务器需要向客户端证明它拥有客户上传的数据，而且这种证明是在不信任云端服务器的情况下完成的。针对这一实际应用的需求，密码学者们提出了两种密码学方法模型。这里简介两个方法的模型：数据可检索性证明(POR)方法[13]、公开可验证的数据持有证明(PDP)方法[14]。

POR 方案由 6 个算法组成，其中 Respond 算法是唯一由证明者 P 执行的算法。其他算法由验证者 V 执行。α 表示 V 执行的状态，假定 α 初始化为空。令 π 表示系统参数的全集。唯一需要明确指出的参数是安全参数 j。⊥表示算法输出失败。

(1) KeyGen[π]→k。算法生成密钥 k。

(2) Encode(F;k,α)[π]→(F_η,η)。算法生成文件的句柄(handle，可视为凭据)η，对给定验证者而言是唯一的。算法同时将 F 转变成(更大)的文件 F_η，并输出一对值(F_η,η)。

(3) Extract(η;k,α)[π]→F。算法 Extract 是一种交互,验证者 V 从 P 那里请求文件。该算法决定了一系列 V 发送给 P 的挑战,以及相应的应答。若成功,算法恢复 F 并输出 F_η。

(4) Challenge(η;k,α)[π]→c。算法输入是密钥 k、文件句柄 η、状态 α 以及系统参数,输出是对文件的挑战 c。

(5) Respond(c,η)→r。算法被 P 使用来生成对 c 的应答。注意,挑战 c 可能来自于 Challenge 或者 Extract。

(6) Verify((r,η);k,α)→b∈{0,1}。算法确定 r 是否是 c 的正确应答。输出 1 表示正确,否则输出 0。

PDP 方案由 4 个算法组成,KeyGen、TagBlock、GenProof、CheckProof,满足:

(1) KeyGen(1^k)→(pk,sk),概率密钥生成算法,输入是安全参数 k,输出一个公钥和私钥对(pk,sk)。

(2) TagBlock(pk,sk,m)→T_m,在客户端运行的算法,生成一个验证元数据,其输入是公钥 pk、私钥 sk 以及一个文件 m,返回验证元数据 T_m。

(3) GenProof(pk,F,chal,Σ)→V,在服务端运行,生成一个 POP(拥有文件的证明)。输入是公钥 pk、一系列文件 F、挑战 chal、对 F 的验证元数据的集合 Σ。输出对 F 中由挑战 chal 确定的 POP(拥有文件的证明)V。

(4) CheckProof(pk,sk,chal,V)→{"success","failure"},在客户端运行,为了验证 POP。输入是公钥 pk、私钥 sk、挑战 chal 以及 POP(拥有文件的证明)V。输出 V 是否是一个正确的对应于 chal 的 POP。

*9.3 计算虚拟化安全

9.3.1 计算虚拟化简介

前面讨论的安全是云计算的"存储"安全,本节讨论云计算的"计算"安全。云计算的一个关键计算就是计算虚拟化技术。虚拟计算(virtualization)技术的引入,打破了真实计算中软件与硬件之间的紧密耦合关系[17]。虚拟计算是相对所谓的"真实计算"而言的,"真实计算"就是将计算建立在真实计算机硬件基础之上。虚拟计算则强调为需要运行的程序或者软件营造一个需要的执行环境,程序和软件的运行不一定独享底层的物理计算资源,对它而言,它只是运行在一个"真实计算"完全相同的执行环境中,而其底层的硬件可能与之前所购置的计算机完全不同。例如,VMware、Xen 等都推出了虚拟化软件。

G. Popek 和 R. Goldberg 认为[18]虚拟计算具有以下 3 个特点:

(1) 保真性(fidelity):强调应用程序在虚拟机上执行,除了时间因素外(会比在物理硬件上执行慢一些),表现为与在物理硬件上相同的执行行为。

(2) 高性能(performance):强调在虚拟执行环境中,应用程序的绝大多数指令能够在虚拟机管理器不干预的情况下,直接在物理硬件上执行。

(3) 安全性(safety):物理硬件应该由虚拟机管理器全权管理,被虚拟出来的执行环

境中的程序(包括操作系统)不得直接访问硬件。

另一个比较广义的定义是:虚拟计算是一种采用软硬件分区、聚合、部分或完全模拟、分时复用等方法来管理计算资源,构造一个或者多个计算环境的技术。

目前已出现不同种类的虚拟化解决方案,由于采用的实现方式和抽象层次不同,使得虚拟化系统呈现不同的特性。计算机系统的设计采用分层结构,通常自底向上分别为硬件、操心系统、程序库、应用程序。其间的接口分别是指令集合(ISA)、系统调用(SysCall)和应用编程接口(API)。理论上,虚拟化技术采用的抽象层次可以在这几层中自由选取。选取的多样性决定了虚拟化技术的多样性。但多样性背后的虚拟化技术实质是一样的:将底层资源进行分区,并向上层提供特定的和多样化的执行环境。下面从虚拟机实现所采用的抽象层次角度对虚拟化系统进行分类[17]:

(1) 指令级虚拟化。通过软件方法,模拟出于实际运行的应用程序(或操作系统)所不同的指令集去执行,采用这种方法构造的虚拟机一般称为模拟器(emulator)。一个典型的计算机系统由处理器、内存、总线、硬盘驱动器、磁盘控制器、定时器、多种 I/O 设备等部件组成。模拟器通过将客户虚拟机发出的所有指令翻译成本地指令集,然后在真实的硬件上执行。例如,Bochs、Crusoe、QEMU、BIRD。

(2) 硬件级虚拟化。硬件抽象层面(Hardware Abstract Layer,HAL)虚拟化实际上与指令集架构虚拟化非常相似,不同之处在于,这种类型的虚拟化所考虑的是一种特殊情况:客户执行环境和主机具有相同指令集合的情况,并充分利用这一个点,让绝大多数客户指令在主机上直接执行,从而大大提高了执行速度。例如,VMware、Virtual PC、Denali、Xen、KVM 等。

(3) 操作系统级虚拟化。一个应用的操作环境包括操作系统、用户函数库、文件系统、环境设置等。如果应用系统所处的这些环境能够保持不变,那么应用程序自身无法分辨出其所在的环境与真实环境之间的差别。操作系统虚拟化技术的关键思想在于,操作系统之上的虚拟层按照每个虚拟机的要求为其生成一个运行在物理机器上的操心系统副本,从而为每个虚拟机提供一个完好的操作环境,并且实现虚拟机及其物理机器的隔离。例如,Jail、Linux 内核模式虚拟化、Ensim。

(4) 编程语言级虚拟化。在应用层次上创建一个和其他类型虚拟机行为方式类似的虚拟机,并支持一种新的自定义的指令集(如 JVM 中的 Java 字节码)。这种类型的虚拟机使用户在运行应用程序的时候就像在真实的物理机器上一样,且不会对系统的安全造成威胁。例如,Java 虚拟机、Microsoft .NET CLI、Parrot 等。

(5) 程序库级虚拟化。在几乎所有的系统中,应用程序的编写都使用由一组用户级库来调用的 API 函数集。这些用户级库的设计能够隐藏操作系统的相关底层细节,从而降低普通程序员的软件开发难度。它们工作在操作系统层上,创造了一个与众不同的虚拟环境,在底层系统上实现了不同的应用程序二进制接口(ABI)和不同的应用程序编程接口(API)。例如 WINE、WABI、LxRun、Visual MainWin。

9.3.2 计算虚拟化的安全

虚拟化系统的安全挑战主要有两个方面:一方面来自计算系统体系结构的改变。虚

拟化计算已从完全的物理隔离方式发展到共享式虚拟化，实现计算系统虚拟化需要在计算性能、系统安全、实现效率等因素之间进行权衡[17]。于是虚拟机监视器和相关具有部分控制功能的虚拟机成为漏洞攻击的首选对象。另外，现有虚拟化系统通常采用自主访问控制方式，难以在保障虚拟机隔离的基础上实现必要的有限共享。另一方面，计算机系统的运行形态发生了变化。虚拟计算允许用户通过操纵文件的方式来创建、复制、存储、读写、共享、移植以及回溯一个机器的运行状态，这些极大地增强了使用的灵活性，却破坏了原有基于线性时间变化系统设定的安全策略、安全协议等的安全性和有效性，包括软件生命周期和数据生命周期所引起的系统安全。

传统计算机的生命周期可以看作一条直线，当前计算机的状态是直线上的一个点。当软件运行、配置改变、安装软件、安全补丁程序时，计算机的状态单调地向前进行。但在一个虚拟计算环境中，计算机的状态更像是一棵树：在任意一点都可能产生多个分支，即任意时刻在这棵树上的任意一点都有可能有一个虚拟机的多个实例在运行。分支是由于虚拟计算环境的可撤销特性与检查点特性所产生的，这使得虚拟机能够回溯到以前的状态或者从某个点重新执行。这种执行模式与一般系统中的补丁管理和维护功能相违背，因为在一般系统中假设系统状态是单调向前进行的。

下面简介一下虚拟机的安全机制。在虚拟机系统中，通常将一些操作系统中的安全及管理函数移到虚拟层里。虚拟层里的核心是高可信的虚拟机监控器。虚拟机监控器通过执行安全策略来保证系统的安全，这首先要求它本身是可信的，本身的完整性可通过专门的安全硬件来进行验证。

虚拟机监控器执行的安全策略对于虚拟系统安全十分重要，例如限制敏感虚拟机的复制；控制虚拟机与底层设备的交互；阻止特定虚拟机被安置在可移动媒体上；限制虚拟机可以驻留的物理主机，在特定时间段限制对含有敏感数据的虚拟机的访问。此外，用户和机器的身份可以用来证明所有权、责任以及机器的历史。追踪诸如机器数据以及它们的使用模式可以帮助评估潜在威胁的影响。而在虚拟层采用加密方式可以用来处理由于虚拟机交换(swapping)、检测点(check pointing)、回溯(rollback)等引起的数据生命周期的问题。

通过虚拟机监控器，多个虚拟机可以共享相同的物理 CPU、内存和 I/O 设备等。它们或者是空间共享的方式，或者是复用的方式使用相同的物理设备，需要通过相应的安全机制保障相互间的有效隔离。虚拟机监控器采用类似虚拟内存保护(虚拟地址访问独立进程地址空间)的方式，为每个虚拟机提供一个虚拟的机器地址空间，然后由虚拟机监控器将虚拟机的机器地址空间映射到实际的机器地址空间中。虚拟机中的操作系统所见的机器地址是由虚拟机监控器提供的虚拟机器地址。

虚拟机监控器运行于最高级别，其次是操作系统。虚拟机监控器具备执行特权指令的能力，并控制虚拟 CPU 向物理 CPU 映射的安全隔离。通过 CPU 硬件的运行级别功能可以有效控制 CPU 虚拟化的安全性。

在程序级虚拟使用环境的安全保障方面，典型的代表就是 Java 安全虚拟机。它提供了包括安全管理器和 Java 类文件认证器等多种安全机制，安全管理器提供在应用程序和特定系统上的安全措施，Java 认证器在.class 文件运行前完成该文件的安全检查，确保

Java 字节码符合 Java 虚拟机规范。针对操作系统虚拟化的安全问题，基于 Windows 操作系统的 Microsoft 虚拟机能阻止恶意用户对 Java Applet 访问 COM 对象、调用 JDBC 等安全漏洞的攻击。

在虚拟机的安全验证方面，典型代表是 ReVirt 系统，该系统采用虚拟机技术提供独立于操作系统的安全验证功能，提供足够信息逐条回放虚拟机上执行的任务，通过建立具有各种依赖关系的攻击事件链，重构出攻击细节，ReVirt 采用了反向观察点和反向断点技术，对虚拟机上的恶意攻击进行检测和回放，提供验证功能。

研究与思考

[1] 云存储安全中的研究对象都是加密后的对象，因此这与一般密码学研究的对象是不同的。于是刺激了诸如密文策略基于属性的加密、可搜索加密、代理重加密、完全同态加密等新的加密方式的研究。思考一下在非可信环境下针对密文操作的特殊加密技术。

[2] 基于云平台的安全工具可能是一个市场需求量很大的产品，你认为什么产品有前途，如何开发。

进一步阅读建议

对密码学感兴趣的读者，可以阅读关于 Fuzzy IBE、ABE、CP-ABE、PRE、FHE 的经典论文，体会密码学原语提出的动机与构造方法的精巧。

[1] Amit Sahai, Brent Waters. Fuzzy Identity-Based Encryption [C]. In Proc. of Eurocrypt05, LNCS 3494, 457-473. 2005.

[2] Vipul Goyal, Omkant Pandey, Amit Sahai, Brent Waters. Attribute-based encryption for fine-grained access control of encrypted data [C], In Proc. of ACM Conference on Computer and Communications Security (CCS06), 89-98. 2006.

[3] John Bethencourt, Amit Sahai, Brent Waters. Ciphertext-Policy Attribute-based Encryption [C]. In Proc. of the IEEE Symposium on Security and Privacy (SP07), 321-334. 2007.

[4] M. G. S. H. Giuseppe Ateniese, Kevin Fu. Improved Proxy Re-encryption Schemes with Applications to Secure Distributed Storage[C]. In Proceedings of the 12th Annual Network and Distributed System Security Symposium (NDSS05). 2005.

[5] Gentry Craig, Halevi Shai. Implementing Gentry's Fully-homomorphic Encryption Scheme[C]. In Proc. of Eurocrypt11, 129-148. 2011.

本章参考文献

[1] NIST Working Definition of Cloud Computing[s]. http://csrc. nist. gov/organizations/fissea/2009-conference/presentations/fissea09-pmell-day3_cloud-computing. pdf.

[2] CSA, Security Guidance for Critical Areas of Focus in Cloud Computing, Version 3[OL]. http://www. cloudsecurityalliance. org/guidance/csaguide. v3. 0. pdf.

[3] 冯登国，张敏，张妍，徐震. 云计算安全研究[J]. 软件学报，2011，22(1)：71-83.

[4] 孙国梓，董宇，李云. 基于 CP-ABE 算法的云存储数据访问控制[J]. 通信学报，2011,(7).
[5] 王海斌，陈少真. 隐藏访问结构的基于属性加密方案[J]. 电子与信息学报，2012,(2).
[6] 单忆南. 基于属性的加密算法[D]. 上海交通大学硕士学位论文，2010.
[7] 杨晓元，王志强，蔡伟艺. 适应性安全的多主密钥 KP-ABE 方案[J]. 中国科学技术大学学报，2011,(7).
[8] 邵俊. 代理重密码的研究[D]. 上海交通大学博士学位论文，2007.
[9] 赵菁，冯登国，杨林，马琳茹. 一个高效的选择密文安全的分类代理重加密方案[J]. 电子学报，2011,(11).
[10] 翁健，陈泯融，杨艳江，邓慧杰，陈克非，鲍丰. 无需随机预言机的自适应攻陷模型下选择密文安全的单向代理重加密方案[J]. 中国科学:信息科学，2010,(2).
[11] 刘艮，蒋天发. 同态加密技术及其在物联网中的应用研究[J]. 信息网络安全，2011,(5).
[12] 汤殿华，祝世雄，曹云飞. 整数上全同态加密方案的重加密技术[J]. 信息安全与通信保密，2012,(1).
[13] Juels A，Kaliski B. PORs：Proofs of retrievability for large files[C]. In Proc. of the 2007 ACM Conf. on Computer and Communications Security (CCS07)，584 - 597. 2007.
[14] Ateniese G，Burns R，Curtmola R. Provable data possession at untrusted stores[C]. In Proc. of the 2007 ACM Conf. on Computer and Communications Security (CCS07)，598 - 609. 2007.
[15] 沈昌祥. 云计算安全[J]. 信息安全与通信保密，2010(12).
[16] 沈昌祥. 云计算安全与等级保护[J]. 信息安全与通信保密，2012,(1).
[17] 金海等. 计算系统虚拟化——原理与应用[M]. 北京：清华大学出版社，2008.
[18] G. Popek，R. Goldberg，Formal Requirements for Virtualizable Third Generation Architectures [J]. Communications of The ACM，17(7)：413-421，1974.
[19] Wei Ren，Linchen Yu，Ren Gao，Feng Xiong，Lightweight and compromise resilient storage outsourcing with distributed secure accessibility in mobile cloud computing[J]，Tsinghua Science and Technology，vol. 16，no. 5，520-528，Oct. 2011.
[20] Wei Ren，Yuliang Liu，A Lightweight Possession Proof Scheme for Outsourced Files in Mobile Cloud Computing based on Chameleon Hash Function，International Journal of Computer Science and Engineering，inderscience，In print，2012.
[21] Wei Ren，Yuliang Liu，A Lightweight Possession Proof scheme for out sourced Files in Mobile cloud computing based on chameleon Hash Function[J]. International Journal of computational science and Engineering. Inderscience. 2012.
[22] 任伟，叶敏，刘宇靓. 云安全的信任管理研究[C]. 2011 年中国计算机学会计算机安全专委会年度会议.

第 3 部分　物联网应用层安全

第 10 章　智能电网安全

在 2009 年美国总统奥巴马提出物联网战略的时候，主要提到了利用物联网解决节能减排、减少碳排放、应对全球气候变暖的智能电网应用。传统电网是以“发电、输电、配电和用电”单一流向的供电模式，据统计这个过程中有 56.2％的能源被浪费，仅在电力生产过程中就有平均 68.8％的能源损失，电力传输和分配过程中还有约 5％的能源损失。因此，提高电力系统效率可带来直接的效益。

根据美国能源部的定义，智能电网[1]是指完全自动化的电力传输网络，能够监视和控制每个用户和电网结点，保证从电厂到终端用户整个输配电过程中所有结点之间的信息和电能的双向流动，其组成部分包括数据采集、数据传输、信息集成、分析优化和信息展现 5 个方面。将物联网技术渗透到发电、输电、配电、用电环节，将提高整个电力系统的效率。

10.1　智能电网概述

10.1.1　智能电网的概念、特征与作用

电力系统是由发电、输电、变电、配电和用电等环节组成的电能生产与消费系统。电力网络是电力系统中除发电设备和用电设备外的部分，主要包括输电、变电和配电 3 个环节。电力网络主要由电力线路、变电所和换流站组成。按功能可分为输电线路、区域电网、联络线和配电网络。

简单地说，电力首先在电站(核能、煤炭、太阳能、地热、风能)生产出来，然后通过层级电力网络，在变电站经过一连串的变压后，将电流从电力源头传输到每个家庭。为了减少远距离传输所产生的线路损耗，一般都是生产和传输高压电(例如我国国家电网就比较推崇较有争议的 1000kV 高压输电)，经过多次变电后输出到目的地。输电线可以架设在空中或铺设在地下。

智能电网是电网的未来发展方向，是以先进的通信技术、传感器技术、信息技术为基础，以电网设备间的信息交互为手段，以实现电网运行的可靠、安全、经济、高效、环境友好和使用安全为目的的先进的现代化电力系统。图 10.1 给出了一个智能电网的基本架构示意图[2]。

国际上智能电网的技术定义为：综合应用现代通讯、计算、控制等技术能够持续不断地适应各种正常操作、运行方式调整的优化运行，并能主动预测和应对电网扰动的电网。因此国际智能电网在功能上希望适应未来数字化信息社会对电能的高可靠性、高质量的要求；适应灵活的发电、用电方式，满足分布式、可再生能源发电接入和灵活的用户供、用电需求；电网具有自适应纠正和自愈能力，主动预防而不是被动地应对紧急情况；持续优化运行以最有效地应用各种资源和设备；电网信息整合更全面；鼓励需求侧响应和用户对

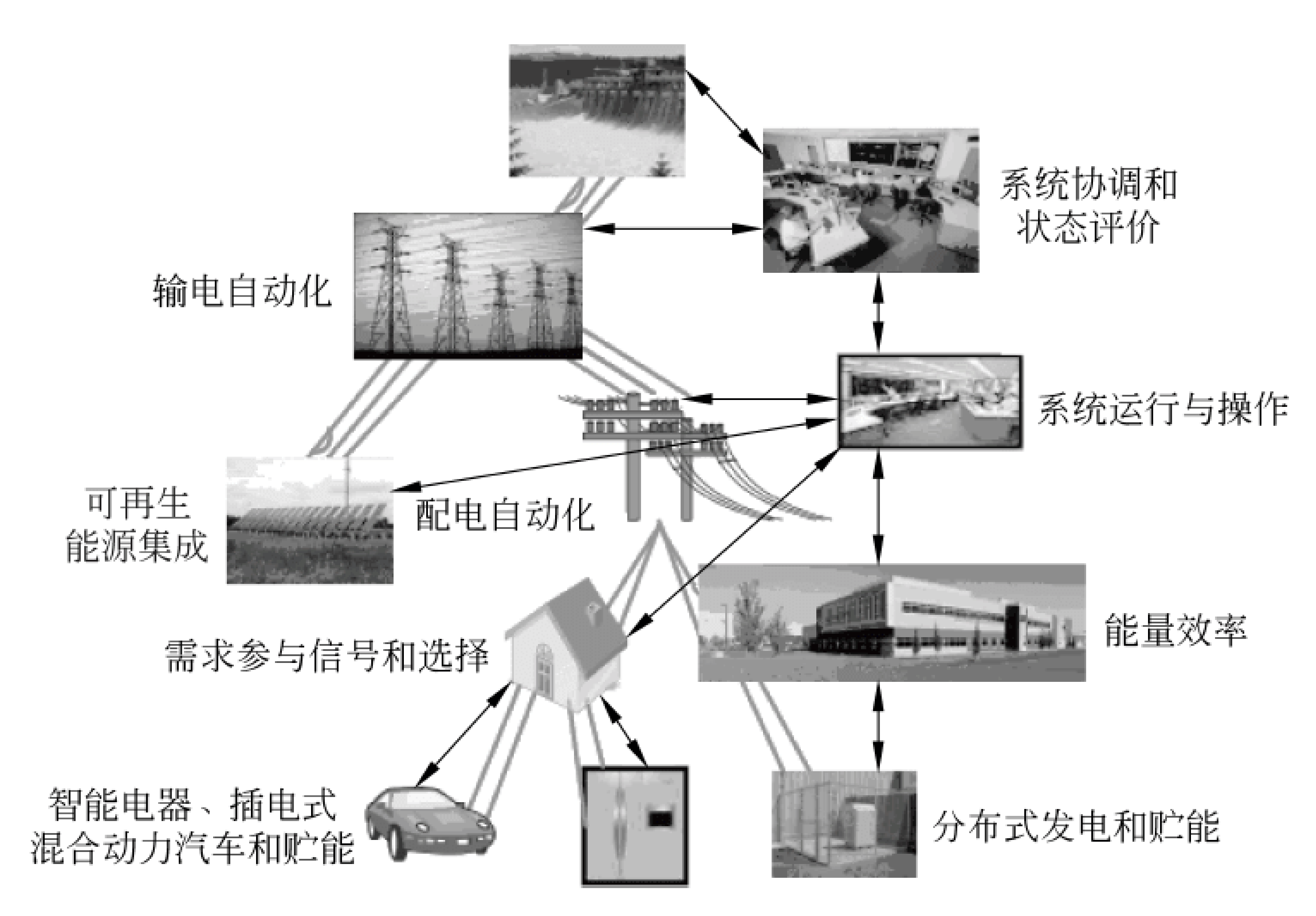


图 10.1　智能电网的基本架构(物理层、通信层以及控制系统)[2]

电网的交互,提供相应的便利接口。总体特点上具有交互性、自愈和自适应、优化能力、预测能力、包容能力、集成能力和更高的安全性。

我国的智能电网概念是指以物理电网为基础(特高压电网为骨干网架、各电压等级电网协调发展的坚强电网为基础),将现代先进的传感测量技术、通信技术、信息技术、计算机技术和控制技术与物理电网高度集成而形成的新型电网。它以充分满足用户对电力的需求和优化资源配置,确保电力供应的安全性、可靠性和经济性,满足环保约束,保证电能质量,适应电力市场化发展等为目的,实现对用户可靠、经济、清洁、互动的电力供应和增值服务。

我国的智能电网具有信息化、数字化、自动化、互动特性,以智能坚强电网为基础,以通信信息平台为支撑,以智能控制为手段,包含电力系统的发电、输电、变电、配电、用电和调度各个环节,覆盖所有电压等级,实现"电力流、信息流、业务流"的高度一体化融合,是坚强可靠、经济高效、清洁环保、透明开放、友好互动的现代电网。

通常认为,智能电网主要具备 9 大特征:

(1) 电网自愈(Self Healing):指在很少或者无需人为干预的情况下,将电网系统的问题元器件隔离,使系统恢复正常运行,并尽量避免用户的供电中断。

(2) 用户交互:可实现和用户的实时双向通信,鼓励并促进用户参与电力系统的运行和管理。

(3) 设备兼容:智能电网可安全、无缝地允许各种不同类型的发电和储能设备接入系统,简化联网过程,实现电网系统中的即插即用。

(4) 质量管理:可依据用户需求对电能质量进行差别化定价,并通过有效的监控手段,检索电能质量的突发事件。

(5) 系统安全:可有效抵御外界对电网的物理攻击或者对网络的攻击,最大限度地降低攻击所造成的损失并快速恢复供电服务。

(6) 信息集成:利用先进的传感、测量技术,智能电网可实现包括监视、控制、维护等各类信息网络系统间的综合集成。

(7) 管理优化:可进一步嵌入能量管理(EMS)、配电管理(DMS)、市场运营(MOS)、企

业资源计划(ERP)等综合管理系统,建设先进的电力系统控制和决策体系,实现管理优化。

(8)资产优化:通过先进的信息和监控技术优化设备和资源的使用效益,提高资产利用效率,并从整体上对网络运行和扩容进行优化,降低运行、维护成本和投资。

(9)市场协调:通过市场化运作提高电力系统的规划、运行和管理水平,利用市场供给和需求的互动,不断推动成本更低的解决方案,促进新技术的开发。

通过物联网技术建设的智能电网,可以提高电力系统效率和可靠性、节能减排、对输变电检测与实时监控、智能高效的配用电管理、高效实时的电力调度、构造坚强安全的电网系统。

特别值得注意的是,除了提升原有电力系统的效率外,智能电网与可再生能源的发展和利用也密不可分。国家电网公司还提出,从 2016—2020 年,我国清洁能源装机比例将极大提高,分布式电源实现即插即用。据预测,到 2020 年,欧洲生产的电力 20%将来自于再生能源。因此,发电方式中出现了所谓分布式发电(Distributed Generation)或者分布式能源(Distributed Energy Resources),电网中分布式可再生能源中产生的电力目前是不可预测的,因此需要加入监测设备进行智能管理(目前我国已跃居成为世界第二大风能市场,还是世界上最大的太阳能光伏设备制造者,众多的电力用户希望把自己产生的绿色电力输送到电网上。由于目前的电网难于调控,这部分电力基本上白白流失了。据不完全统计,全国各主要风力发电企业 2009 年丢失的电量超过 15 亿 kW·h,约占风电总发电量的 12%)。同时,电动汽车(插拔式混合动力电动汽车 PHEV)也给智能电网带来了挑战,即需要一个智能机制使得电力消耗平稳进行,并根据电力生产的情况适时地将电力从电网中引出来。值得注意的是,可以用数百万的汽车电池作为缓冲来平抑电力峰值波动,这便是汽车电力储能 V2G(Vehicle to Grid,车辆到电网)的理念。由于大部分车辆 95%的时间是处于停驶状态,车载电池可以作为分布式储能单元。据估算,每辆车可为电力公司带来 4000 美元的价值。

这里我们大胆地进行一个设想:2020 年每个家庭均拥有太阳能或风能发电设备,至少拥有两组蓄电池,一组用于蓄电,一组用于供电,隔天切换。供电的电能可用于家庭的用电,多余的电能可以保存到自己家的 PHEV 中,PHEV 可作为家庭的移动电源和备份电源,如果还有剩余电能,甚至可以供给其他家庭使用。在输电价格高昂时,多个家庭的 PHEV 可将保存的多余电量输入电网,用于小社区的供电。

智能电网中使用的智能电表将改变家庭的用电习惯,提高用电效率。在我国智能电表仅能单向作业,可满足远程抄表(Automatic Meter Reading,AMR),属于智能电表智能化程度最低的一级,主要用于替换传统机械表。高级计量体系(Advanced Metering Infrastructure,AMI)可以实现中心 SCADA 与智能电表间的双向通信,内容包括动态定价、需求响应(Demand Response)、负荷曲线,警报以及自动计费状态等。另外随着绿色能源逐步走向分布式发展,智能电表还将满足大量的绿色能源净计量(Net Metering)。此外,智能电表还可在任何地点对家庭和公共用电设备进行管理控制(例如家庭能源管理),提供多种增值服务。

因此,智能电网的更广义的作用是解决可再生能源大规模并网问题;推动未来电动汽车、智能设备、智能家电、智能建筑、智能交通、智能城市等智能经济形态广泛应用;为客户提供实时电价和用电信息,引导客户合理用电,提高能源利用率。智能电网的网络架构与安全手段对于很多其他领域同样可以参考借鉴。

10.1.2 智能电网的通信与网络架构

2009 年 9 月，美国国家标准与技术研究所（NIST）提出了关于智能电网互操作标准的框架与路线图，明确了推进标准化工作的 8 个优先发展领域：广域网状态可感知、需求响应、电能存储、电力交通、网络安全、网络通信、先进的计量基础设施和配网管理[3]。其中，与网络技术直接相关的有 3 个领域：

（1）网络通信（Network Communication）。要求针对智能电网各个关键领域的应用和操作器的网络通信需求，实施和维护合适的安全和访问控制手段。该领域覆盖电力专网和公共网络。这一部分相当于描述了智能电网中物联网网络层的功能。

（2）先进的计量基础设施（AMI）。能够提供双向通信，既能为多个功能系统使用，也能使授权的第三方与用户设备和系统交换信息，AMI 系统能为用户提供透明的实时电价感知功能，也能帮助供电方实现必要的减负目标。这一部分相对于描述了智能电网中物联网感知层的功能。

（3）网络安全（Cyber Security）。这将在下一节重点介绍这一点，在介绍网络安全之前，首先需要明确通信与网络架构。

NIST 提出了智能电网的概念参考模型[3]（见图 10.2），将智能电网划分为 7 个领域：输电（transmission），配电（distribution），运行和操作（operations），大容量发电（bulk generation），电力市场（markets），电力用户（customers）以及电力供应（service provider）。HAN 是家庭域网络（Home Area Network），BAN 是建筑域网络（Building Area Network），IAN 表示工业域网络（Industrial Area Network）。

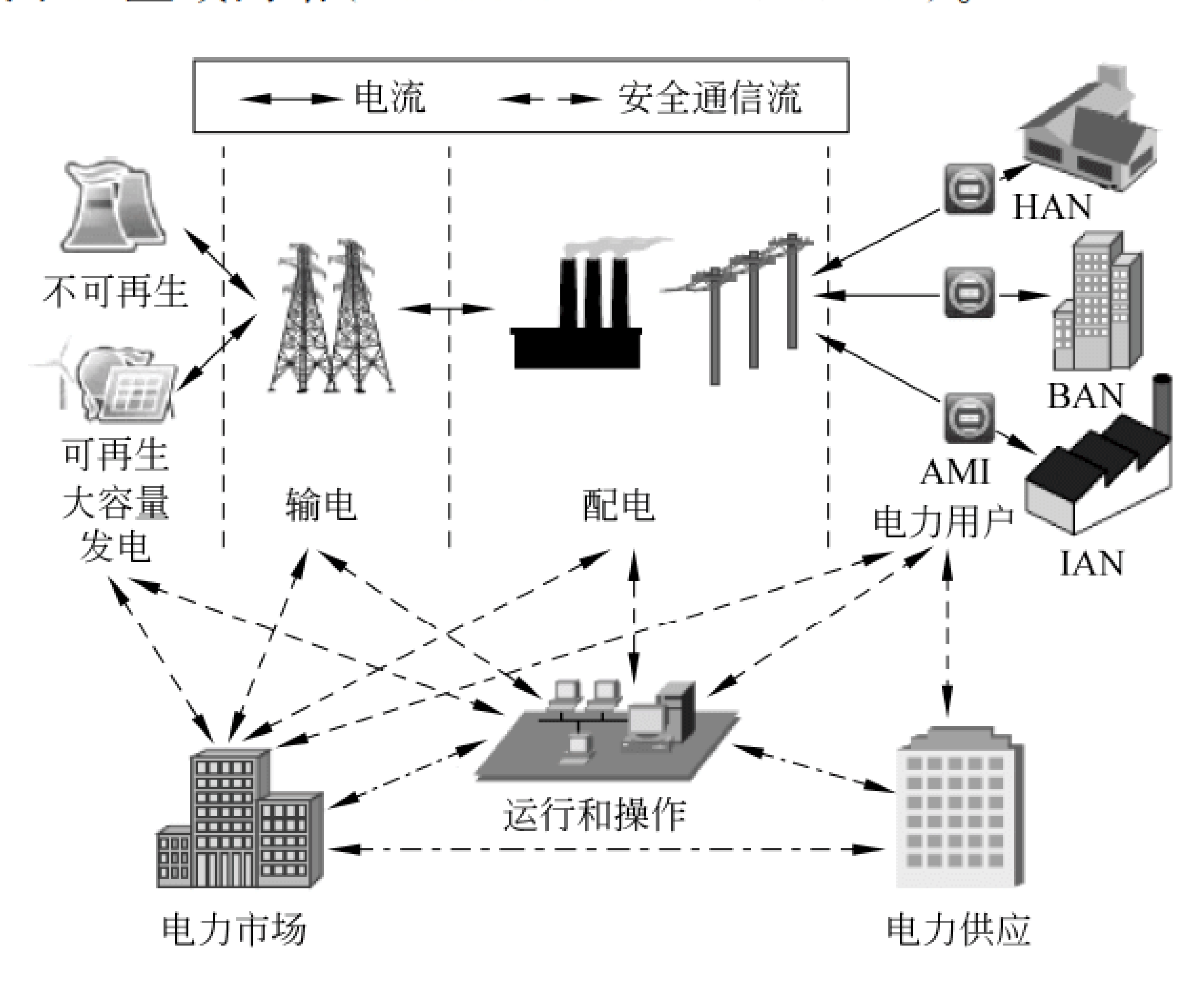


图 10.2　NIST 智能电网的概念参考模型[3,11]

其中，供电部门为终端用户提供供电服务；用户不仅是电力系统的终端用户，也能够参与发电、输电和管理电能的使用。主要分为 3 类：居民用户、商业用户、工业用户。大容量发电单位既能发电也可储电。这 7 个领域覆盖电力行业的各个环节，每个领域和子领域中的执行单元（软件、硬件设备和系统）通常需要通过网络与其他域的执行单元进行交互。每个领域的具体内容见图 10.3。

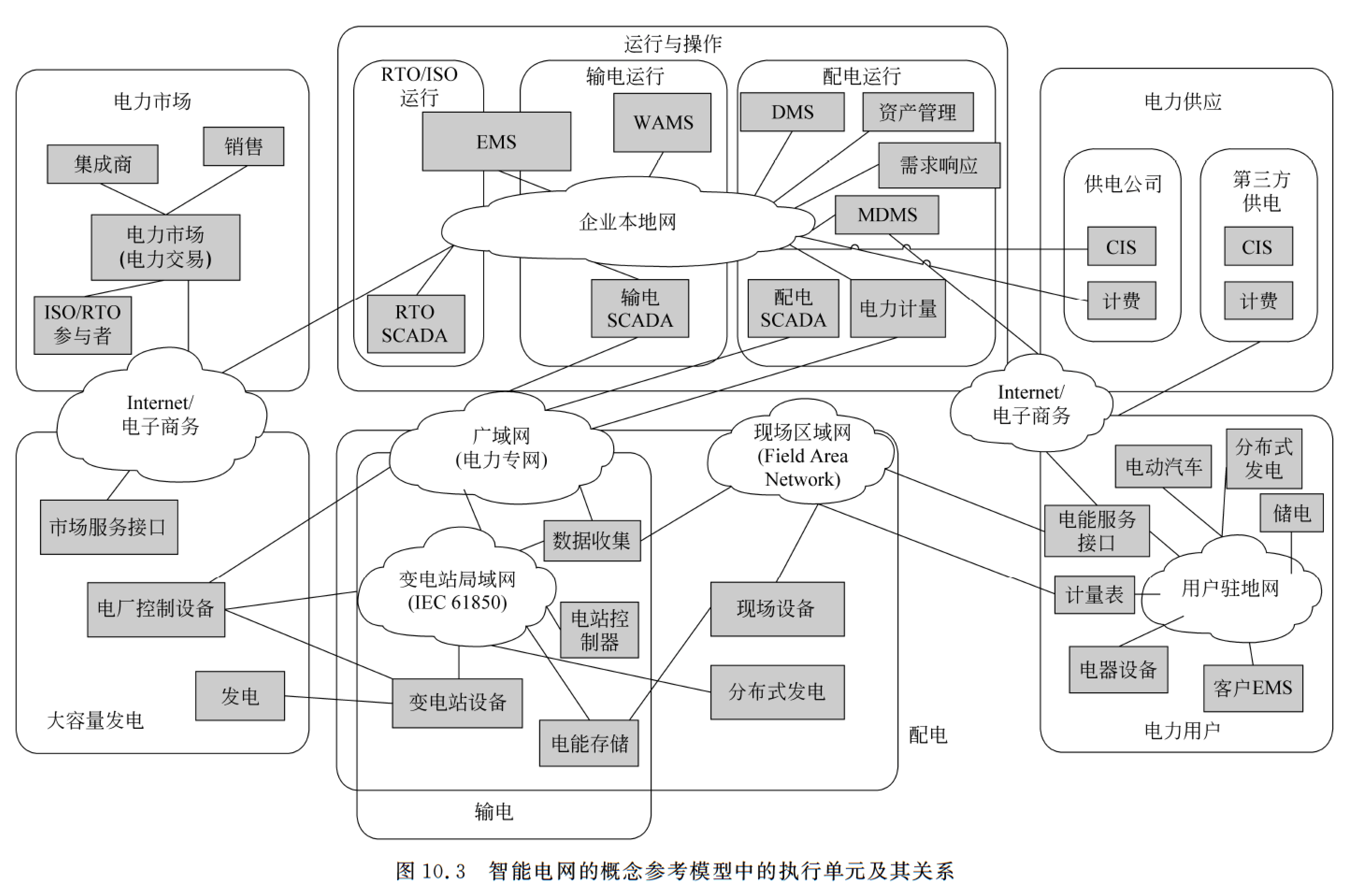


图 10.3　智能电网的概念参考模型中的执行单元及其关系
（注：ISO/RTO 为 Independent System Operator/Regional Transmission Organization 缩写）

文献[4]给出了智能电网中网络技术架构的详细说明，见表 10.1。其实，在 HAN 接入方面还可以采取 M2M 的无线通信远程接入方式，因此这里作了些改动。

表 10.1 智能电网的网络技术架构

网络成分	可采用的网络技术	应用说明
广域网（WAN）	IP、DWDM	提供电力数据网（骨干网）、因特网的网络互联和路由等功能
	MPLS、MPLS VPN	骨干网中提供标记交换，隔离不同业务的流量
	ATM	保护原有投资技术，已逐渐退出应用
接入网（AN）	SDH	为接入城域网、广域网提供物理通道
	MSTP	可用于 LAN（以太网）接入城域网
	GPRS	以无线方式接入广域网
	PON	无源光网络，提供光纤接入方式
企业本地网（LAN）	IEEE 802.3、802.1d、802.1q	电力企业的 Intranet
现场区域网（FAN）	RS-485、PROFIBUS 等传统的现场总线	电厂、变电站等生产控制领域
	工业以太网（802.3、802.1d、802.1q）	电厂、变电站等生产控制领域 IED 设备互连
	N-PLC、B-PLC、BPL（窄带、宽带电力线载波通信）	用于计量、仪表数据采集等数据的传输
	无线传感器网络（802.15.x）	输电、配电、用电侧的数据采集、监测和监控
	物联网、RFID	设备巡检中标签数据的采集
用户驻地网及家庭网络（HAN）	PON、EPON、FTTH	智能化住宅小区，提供家庭用户的光纤接入
	N-PLC、B-PLC、BPL	提供驻地网及用户家庭网络接入，远程抄表、因特网接入
	无线局域网 802.11、M2M GPRS 模块	用户驻地网或家庭网络接入、远程抄表
	无线传感器网络（802.15.x）	用于 HAN 中智能家居，家电控制

电力行业的专用骨干网是电力数据网络和电力信息网络，主要采用 MPLS 技术，利用 MPLS VPN 对不同业务之间进行逻辑隔离，通过为不同的业务系统划分虚拟专用网，有效隔离不同业务，保证业务的安全和可靠运行。另外，MPLS 可针对不同业务需求提供服务质量保证。骨干网可在 VPN 的部属策略和流量工程的规划两个方面进行技术的深化应用。

分布式传感器网络解决从电力系统远程监测、状态感知、远程控制，到用户侧的实时计量和智能家居交互的需求。可采用的方法是需要 IP 层的结点或者无需 IP 层的结点，前者采用 6LoWPAN 技术，这在第 8 章已经介绍了。后者将应用层直接映射到数据链路

层，典型应用如 IEC61850 中定义的具有低延时要求(1～4ms)的变电站事件通用对象 GOOSE(Generic Object Oriented Substation Events)报文。

10.2　智能电网安全

10.2.1　智能电网的安全架构与安全需求

文献[5]从全局上讨论了智能电网的完整信息安全体系。智能电网的信息安全需求总体上包括面向电网核心应用及业务系统的安全防御、电网基础信息网络和重要信息系统的安全保障、复杂大系统下的主动实时防护、安全存储、网络病毒入侵与防范、恶意攻击的防范、网络信任体系的建立、新的密码技术等。

由于安全技术种类繁多，需要建立一个防护对象层次化、防护结构合理化、技术管控动态化的安全技术体系防御架构，以便依照智能电网的不同发展阶段，满足信息安全建设的相应安全需求。安全技术上要强调系统平台深度集成、安全信息高度共享、管控流程高度贯通、业务安全高度互动，涵盖电网核心应用安全各个环节；统一建设和总体规划智能电网信息安全技术体系架构。安全技术体系建设的总体架构如图 10.4 所示。



图 10.4　智能电网信息安全技术体系[5]

安全技术体系建设总体架构的主要内容包括技术安全管控、物理数据安全、主机终端安全、网络通信安全、业务应用安全和安全支撑平台等。

技术安全管控保证安全技术体系的可控、能控和在控，满足安全技术建设的全局监测

和集中管控。它研究信息系统的安全需求、业务系统的安全机制与安全模型,进行全局代码安全性测试与安全技术全生命周期管控,逐步将核心安全技术的可靠性、稳定性纳入信息安全技术建设体系,在安全技术全生命周期各个阶段确保技术系统建设的安全性和有效性,有效提高核心安全技术质量,保障安全技术体系的深度防御建设。

物理数据安全包括软硬件双向安全保护。主机终端安全实现各种计算机终端设备的安全防御,范围涵盖服务器、终端计算机、智能移动接入设备和智能电表等。网络通信安全的主要目标是实现重要网络和信息通信的主动可控、数据传输加密等技术,有效地对网络边界数据流的通信交换进行安全检测和访问控制,防范内网重要数据非法外泄和外部网络有害信息流入,实时中断、调整或隔离不正常或具有破坏性的网络数据传输或通信行为,及时识别入侵攻击模式,提升智能电网网络通信的可控性、可信性和稳定性。业务应用安全主要保证智能电网业务应用的连续性、完整性和可靠性,以及工作流程的稳定性和应用操作的实时性。

安全支撑平台为安全技术体系提供信息安全基础设施保证,主要目标是构建一个智能电网可相互信任的信息交互环境,为其他安全技术建设提供安全认证标识和决策依据。其中安全统一认证管理提供准确有效的身份认证和访问控制。综合安全审计和安全风险动态评估可实现安全技术体系运行环境和重要信息业务系统安全性的实时监控评测,对体系建设和系统安全现状进行有效的风险预警、安全性能评测分析、动态安全风险管理与体系建设完善升级。安全技术标准建设用于构建一套完整的安全技术体系策略执行标准,实现信息安全技术体系的一体化建设,提升智能电网安全技术集约管控能力和整体标准化水平。

在具体的安全需求方面,NIST 2010 年 8 月发布了智能电网安全指南(Guidelines for Smart Grid Cyber Security)[6],文档中分析了安全的边界,并根据安全的边界进行了安全需求分析。安全的边界采取逐步求精的方法,分析了可能涉及的实体和交互。这其实示范了一种分析安全需求的方法。首先划分领域,也就是一群具有相似目标的组织、建筑、个人、系统、设备或执行单元(actor)等。上一节曾经提到,根据 NIST 智能电网的架构与路报(framework and roadmap)文档,智能电网的 7 个领域之间的交互关系见图 10.5。

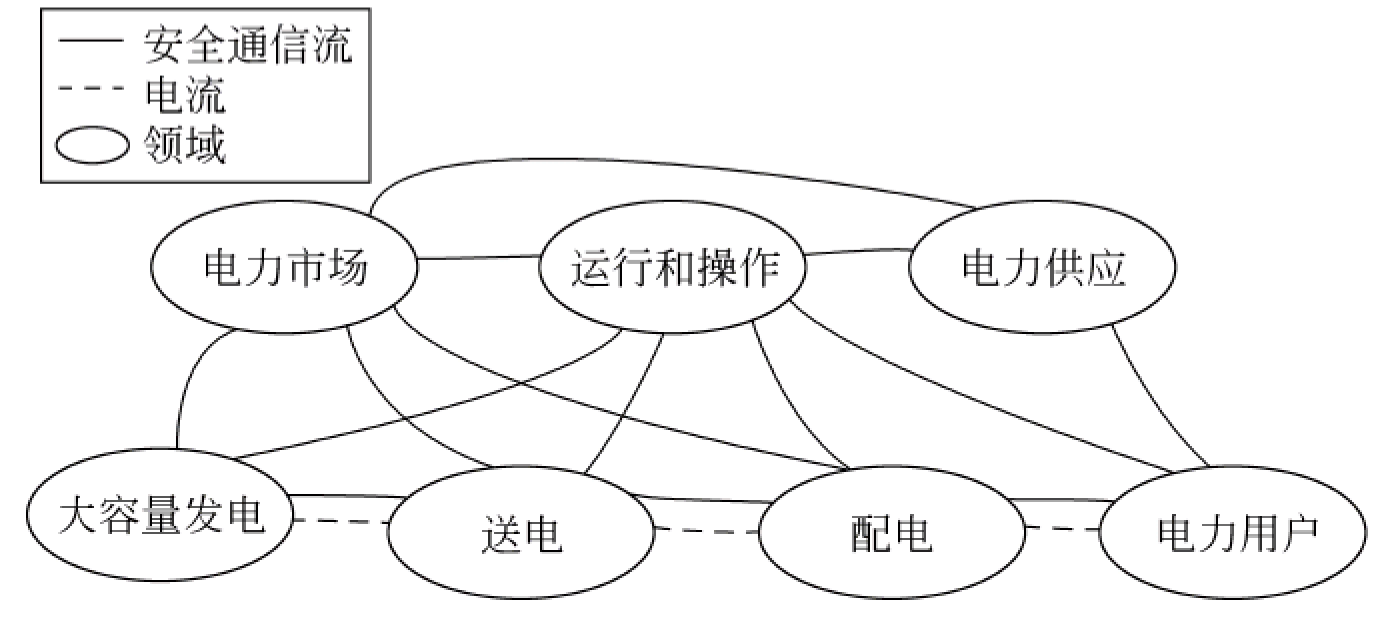


图 10.5　智能电网各个域之间交互的安全通信流和电流

在 7 个领域中,有大量地获取、传输、保存、编辑和处理智能电网应用信息的行为。域之间会有一些交互,但是同一个域不必要具有相似的安全需求,且某些域可能包含其他域的组件。将每个域继续细化(如图 10.6,译自文献[6]),可以得到 130 个可能的接口,这

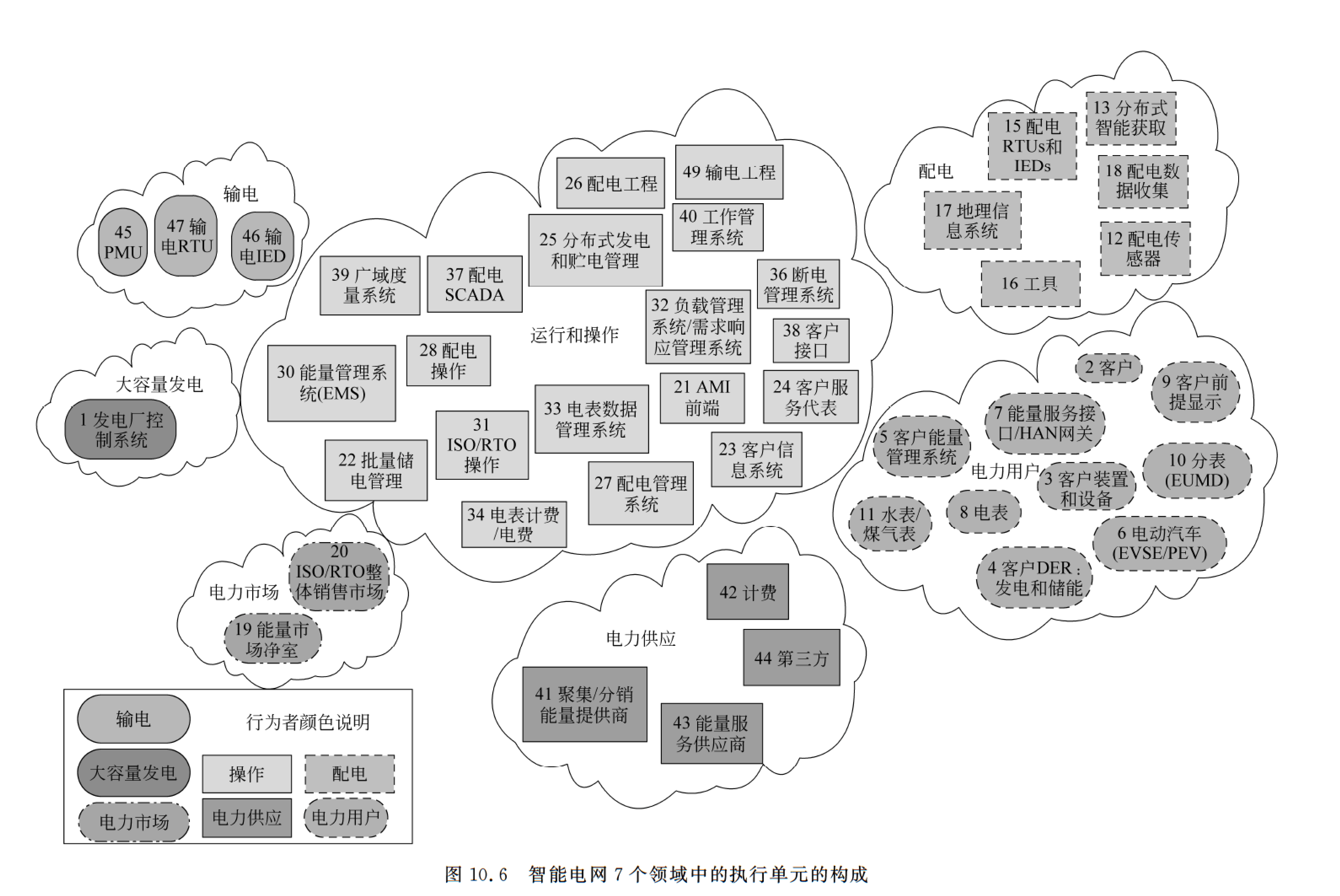


图 10.6　智能电网 7 个领域中的执行单元的构成

些接口分配给22个具有类似安全特征的大类中。对于每个大类，将评估设备失效、入侵以及其他安全对智能电网的系的性能、信息以及信息系统造成的威胁，包括3个方面：破坏机密性，即对信息的非授权泄漏；破坏完整性，即对信息的非授权修改和破坏；破坏可用性，即失去对信息和信息系统的使用和访问。NIST通过该文档给出了所有行为者之间涉及的交互的详细的安全需求说明。

文献[7]认为智能电网面临的主要风险如下：

(1) 网络通信更复杂，无线局域网、移动通信网络、卫星通信、智能传感网等多种通信方式、多种网络协议并存，使得电网通信网络更加复杂。信息在传输过程中存在被非法窃听、篡改和破坏的风险。

(2) 信息交互更频繁；信息系统集成度、融合度更高，系统依赖性更强，业务系统之间、业务系统与外界用户之间实时交互更加丰富与频繁。同时，海量交互信息有可能导致数据吞吐量过大，造成网络波动、业务过载。终端用户交互信息存在泄漏、篡改和破坏的风险。

(3) 新技术应用更广泛：随着新型无线通信技术、智能设备、虚拟化、物联网、云计算、多网融合等前沿技术逐步应用、发展和成熟，各类信息安全问题可能凸显。

(4) 智能终端接入更多：智能表计、智能家电、分布式能源设备等多种智能终端大量接入，业务终端数量庞大、类型多样，存在信息泄漏、非法接入、被控制的风险。

10.2.2 智能电网的安全问题简介

近年来，智能电网安全在美国已经成为热门研究问题，但大多数研究还是围绕在提炼和明确研究的问题方面。虽有一些研究成果，但还需要更长的时间来检验。虽然可以利用很多通信安全、网络安全、控制系统安全的现有方法去解决一些问题，但我们认为，智能电网应还有一些特色的具体安全问题，这些问题应该从智能电网的应用特点出发去寻找。

文献[8]指出了企业网络安全与智能电网安全之间的区别。企业网络安全中的防火墙、入侵检测系统、虚拟专网(VPN)等手段在控制和自动化网络安全方面是不够的。企业网与智能电网安全的区别有3个方面：

(1) 不同的安全目标。企业网的安全目标是保护数据。因此数据完整性、数据保密性、数据可用性是3个主要关注的地方。但在智能电网方面，首先是保障人员安全，其次是保障系统可靠性，如避免断电(outage)或电压过低(brownout)等，最后就是保护设备和电力线。

(2) 不同的安全架构。企业网中的数据服务器在网络的中心，与边界结点相比需要更多的保护。在智能电网中，企业管理系统(Enterprise Management Systems，EMS)居于中心位置，但远程终端单元(Remote Terminal Units，RTU)和电力线通信(PLC)处在边界位置。通常只有直接被RTU/PLC控制的设备才会对人身安全产生危害。EMS/SCADA以及数据日志服务器不会直接造成人身危害。因此，在智能电网中边界结点需要和中心设备同级别的安全保护。

(3) 不同的技术基础。在企业网中，操作系统通常是Windows、UNIX和Linux，以太网以及基于IP的协议连接所有的设备。因此可以使用基于这些共同架构的统一的安

全解决方案，但是在智能电网通信系统中，除了上述操作系统外，还有很多专属操作系统和网络设备，具有不同的通信协议（如 IEC61850、DNP3.0、ICCP 等），而不仅仅是 TCP/IP 协议栈。因此，开发统一的基于主机的或者基于网络的安全解决方案是非常困难的。

文献[9]指出了智能电网安全领域研究面临的挑战，其中比较有特色的包括：

（1）AMI 的安全性。地理上分布的架构以及高可用性的需求使得安全和隐私问题得到更多的关注，对于 AMI 而言，需要能够远程证明（remote attestation）AMI 模块，能够检测到对硬件的破坏，阻止电表被操纵；需要研究因共性安全失败导致的问题，如恶意代码的传播、远程漏洞利用等；基于模型的异常检测方法，可根据已知攻击模式确定攻击，研究攻击检测算法；安全和隐私的权衡，包括对消费习惯的推理、匿名机制、静态数据和移动数据的匿名性等。

（2）信任管理。智能电网的动态性要求信任管理来评价系统输入、输出的可信度。动态的信任分配，要考虑到威胁和可能的安全失败的情况（如暴露的认证者、没有打补丁的系统），也要考虑到电网紧急情况，如连锁故障（cascading failures）、自然灾害、人为事故等。基于数据源（如 SCADA 设备、相关器件等）的信任管理，对低信任系统（物理上没有保护的、资源受限的系统）的验证、信任验证的机制或算法以及信任操纵机制的影响分析。随着数据和验证源的增加（如更多的传感器），以及在整个 AMI 中信任需求的积聚，导致信任的聚集。

文献[10]指出智能电网可靠性比传统电网面临更多的挑战，这是因为：

（1）电量的不确定性、多样性、和因为考虑到环境的可持续性导致的电力供应的分布性，这些使得实时的电量模式与离线设计和分析的模式有明显的不同。

（2）市场因素加重了更多的远距离提高电压输电，可靠电力的范围在减小。

（3）在远端的电力运行越来越频繁，因为“不充分的”投资和有限的措施，电力消费和峰值需求的增长，基础设施的老化，最大化对监控、分析和控制投资利用率。

（4）电力运行的实体具有更大的“动静”，更多复杂的问题需要更短的决策时间和更少的错误。

（5）大量的利用分布式的资源模糊了输电和配电之间的界限，电力系统的复杂性和易变性被强化了。

文献[11]在智能电网的隐私保护方面给出了较为详细的说明。根据 NIST 的调研，有 4 种典型情形需要考虑到隐私保护：

（1）欺诈的情况，特别是在不同的地点消费能量的情况，如插电式混合动力汽车（Plugin Hybrid Electric Vehicle，PHEV）。电表系统不能允许任何操作人员滥用或修改收集的数据。NIST 的报告中分析了 2 个例子，一个是当房东和租户都拥有 PHEV 时，他们需要分别收费。为保护租户的隐私，在智能电表和 PHEV 间的通信要通过安全连线和能力服务通信接口（Energy Services Communication Interface，ESCI）来认证，这一接口由供电装置（utility）或者车辆制造商提供。另一个例子是 PHEV 需要有一个通用的登记和准入过程。NIST 认为供电装置应该提供诸如准入、登记、初始化连接、在供电装置和 PHEV 重复建立连接、PHEV 的收费状态信息、正确的收费单。

（2）HAN 智能电表中的数据可能泄漏户主电器使用的行为。而且，这些数据可用于

追溯特定的时间和位置的能量消费情况，从而分析出使用电器的情况和行为。例如，电器供应商可能想知道使用者是如何以及为何使用他们的产品。这些信息影响了电器的保修期。同时这些数据可以推导出目标市场。

(3) 获得接近实时的能量消费数据可能推导出居住情况，人们是如何居住的，在干什么等。对于这一点已经有一些研究成果。

(4) 个人生活方式信息可以从能量消费数据中推导出，这对于某些第三方是有用的信息。

NIST 发布了一个关于电力用户隐私影响评估(Privacy Impact Assessment，PIA)的报告，建议 10 个设计原则来应对智能电网隐私保护：

(1) 一个组织应该确信信息安全和隐私策略与行为是存在的，且具有明文规定并遵守。审计功能应可以监控所有的数据访问和修改。

(2) 在收集和共享个人信息和能量使用数据之前，应该发布一个陈述清晰的注意事项。

(3) 可用的措施应该呈现给所有的用户。在收集、使用、透露他们的个人信息时，组织需要获得用户的同意。

(4) 只能从个人那里收集最少的需要完成所声称的目的的个人信息。对这些信息的使用应该遵照隐私保护原则。

(5) 信息只能用于在收集时所确定的目的，且只能透露给授权方。个人信息应该尽可能地聚集或者匿名化，以减少记载的计算机记录。个人信息应该只保存它需要完成既定目的的那段时间。

(6) 组织应允许个人检查他们的个人信息，并可请求修改其中不准确的地方。

(7) 个人信息数据仅仅用于收集时的既定目的，透露给非与服务接受者明确达成共识的第三方是不允许的。

(8) 个人信息的任何形式不应被非授权地修改、复制、公开、访问、使用、丢失或者窃取。

(9) 组织应该确保数据使用信息是完整的、精确的、与注意事项中声明的目的相关。

(10) 隐私策略应该对服务接收方公开。这些服务接收方应能够质询组织者是否遵循了这些隐私保护规则。

研究与思考

[1] 智能电网安全中的健壮性是一个关键的需求，因此安全机制的健壮性也很重要，如何保障安全机制的健壮性。

[2] 智能电网中在收集用电信息的时候，详细的用电过程可能导致推导出用户的生活行为，这将会导致用户隐私的泄漏，如何解决这一问题。

[3] 家庭智能电源管理与新能源汽车，特别是 PHEV 是未来的新应用，设想一下这些应用中会遇到哪些安全问题。

[4] 分布式发电与储能数据以及电力消费数据最终都要上传到 SCADA，才能形成高效的电力配电。

如何保障这些数据的安全性。

[5] 思考 V2G 应用模式是否存在安全问题。

进一步阅读建议

[1] A. Banerjee，K. K. Venkatasubramanian，T. Mukherjee，S. K. S. Gupta，Ensuring Safety，Security，and Sustainability of Mission-Critical Cyber-Physical Systems[J]，Proceedings of the IEEE，vol. 100，no. 1，pp. 283-299，Jan. 2012.

[2] Xudong Wang，Ping Yi，Security Framework for Wireless Communications in Smart Distribution Grid[J]，IEEE Transactions on Smart Grid，vol. 2，no. 4，pp. 809-818，Dec. 2011.

[3] M. M. Fouda，Z. M. Fadlullah，N. Kato，Rongxing Lu，Xuemin Shen，A Lightweight Message Authentication Scheme for Smart Grid Communications[J]，IEEE Transactions on Smart Grid，vol. 2，no. 4，pp. 675-685，Dec. 2011.

[4] Dapeng Wu，Chi Zhou，Fault-Tolerant and Scalable Key Management for Smart Grid[J]，IEEE Transactions on Smart Grid，vol. 2，no. 2，pp. 375-381，June 2011.

[5] A.-H. Mohsenian-Rad，A. Leon-Garcia，Distributed Internet-Based Load Altering Attacks Against Smart Power Grids[J]，IEEE Transactions on Smart Grid，vol. 2，no. 4，pp. 667-674，Dec. 2011.

[6] Y. Yan，Y. Qian，H. Sharif，D. Tipper，A Survey on Cyber Security for Smart Grid Communications[J]，IEEE Communications Surveys & Tutorials，vol. PP，no. 99，pp. 1-13，2012.

本章参考文献

[1] NIST Smartgrid，http://www.nist.gov/smartgrid/.

[2] H. Khurana，M. Hadley，Ning Lu，D. A. Frincke，Smart-Grid Security Issues[J]，IEEE Security & Privacy，vol. 8，no. 1，pp. 81-85，Jan.-Feb. 2010.

[3] NIST Framework and Roadmap for Smart Grid Interoperability Standards，Release 1.0 (NIST SP 1108) [S]，Sept. 2009. http://www.nist.gov/public_affairs/releases/upload/smartgrid_interoperability_final.pdf.

[4] 徐磊.智能电网的网络通信架构及关键技术[J]. 电气技术，2010 年 08 期，16-20.

[5] 关志涛，颜立，何杰涛，丁涛. 面向智能电网的信息安全技术展望[J]. 陕西电力，2010 年 06 期.

[6] NIST IR 7628，Aug. 2010，Guidelines for Smart Grid Cyber Security [S] http://csrc.nist.gov/publications/nistir/ir7628/introduction-to-nistir-7628.pdf.

[7] 余勇，林为民，邓松，车建华. 智能电网中的云计算应用及安全研究[J]. 信息网络安全，2011 年第 06 期，41-43.

[8] Y. Yan，Y. Qian，H. Sharif，D. Tipper，A Survey on Cyber Security for Smart Grid Communications[J]，IEEE Communications Surveys & Tutorials，vol. PP，no. 99，pp. 1-13，2012.

[9] Y. Mo，T. H.-J. Kim，K. Brancik，D. Dickinson，Heejo Lee；A. Perrig，B. Sinopoli，Cyber-Physical Security of a Smart Grid Infrastructure[J]，Proceedings of the IEEE，vol. 100，no. 1，pp. 195-209，Jan. 2012.

[10] K. Moslehi, R. Kumar, A Reliability Perspective of the Smart Grid[J], IEEE Transactions on Smart Grid, vol. 1, no. 1, pp. 57-64, June 2010.

[11] J. Liu, Y. Xiao, S. Li, W. Liang, C. Chen, Cyber Security and Privacy Issues in Smart Grids[J], IEEE Communications Surveys & Tutorials, vol. PP, no. 99, 1-17, 2012.

[12] Wei Ren, Jun Song, Yu Yang, Yi Ren, Lightweight Privacy-aware yet Accountable Secure Scheme for SM-SGCC communications in smart grid[J], Tsinghua Science and Technology, vol. 16, no. 6, pp. 640-647, Dec. 2011.

[13] Wei Ren, Jun Song, Min Lei, Yi Ren, BVS: A Lightweight Forward and Backward Secure Scheme for PMU Communications in Smart Grid[J], International Journal of Digital Multimedia Broadcasting, vol. 2011, Article ID 382147, 9 pages, 2011.

第 11 章　EPCglobal 网络安全

RFID 系统在物联网背景下演变成连接到 Internet，可访问开放和海量标签（EPC 编码）数据库的 EPCglobal 网络（EPCglobal Network）。EPCglobal 网络是 RFID 在物联网应用中的典型网络，它可以广泛应用于需要利用标识（Electronic Product Code，EPC）进行物体追踪的场景中，如物流配送和全球供应链管理、精准农业等。EPCglobal 网络是物联网的一个典型网络架构（第 1 章曾经提到，另外两种网络架构是基于 6LoWPAN，即 IPv6 over IEEE 802.15.4 的智能物体架构，以及基于广域无线移动通信网络的 M2M 架构）。

11.1　EPCglobal 网络概述

11.1.1　EPCglobal 网络简介

EPCglobal 网络是实现自动即时识别和供应链信息共享的网络平台。通过 EPCglobal 网络，提高供应链上贸易单元信息的透明度与可视性，以此各机构组织将会更有效运行。通过整合现有信息系统和技术，EPCglobal 网络将提供对全球供应链上贸易单元即时准确自动地识别和跟踪。

EPCglobal 是一个中立的、非赢利性标准化组织，由国际物品编码协会（European Article Numbering Associtation，EAN）和美国统一代码委员会（Uniform Code Council，UCC）两大标准化组织联合成立。EPCglobal 的主要职责是在全球范围内对各个行业建立和维护 EPC 网络，保证供应链各环节信息的自动、实时识别采用全球统一标准。通过发展和管理 EPC 网络标准来提高供应链上贸易单元信息的透明度与可视性，以此来提高全球供应链的运作效率。

EPCgloabal 起初由美国麻省理工学院（MIT）的 Auto-ID 中心提出，2003 年 10 月后，Auto-ID 中心研究功能并入 Auto-ID 实验室，总部设在 MIT，与其他 5 所大学通力合作研究和开发 EPCglobal 网络及其应用。（这 5 所大学是英国剑桥大学、澳大利亚阿德莱德大学、日本庆应大学、中国复旦大学和瑞士圣加仑大学）。

EPCglobal China 是经国际物品编码中心授权的、EPCglobal 在中国的唯一代理，负责我国 EPC 的注册、管理及标准化工作，提供相关技术支持，促进 EPC 系统在我国的推广和实施。

11.1.2　EPCglobal 物联网的网络架构

典型的 EPC 物联网由信息采集系统、PML 信息服务器、对象名解析服务器（ONS）和

Savant 系统 4 部分组成，如图 11.1 所示。这一部分引自文献[4]，下面分别加以介绍。

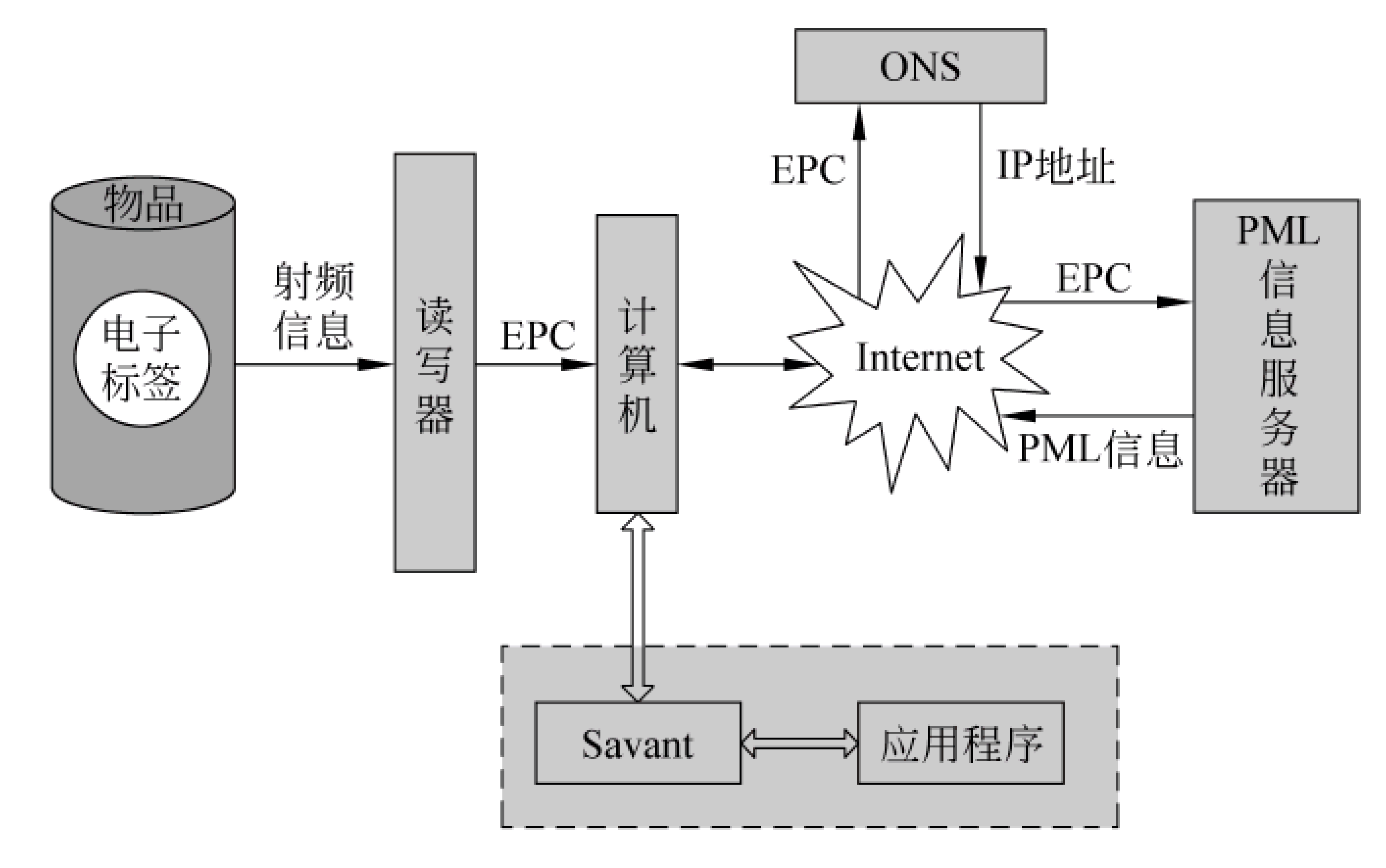


图 11.1　物联网的系统组成

1. 信息采集系统

信息采集系统由产品电子标签、读写器、驻留有信息采集软件的上位机组成，主要完成产品的识别和 EPC 的采集与处理。

EPC 是 Auto-ID[1] 中心为每个物理目标分配的唯一可查询的标识码，其内容为一串数字，可代表产品类别和制造商、生产日期和地点、有效日期、运往何地等信息。同时，随着产品在工厂内的转移或变化，这些数据可以实时更新。通常，EPC 可存入硅芯片做成的电子标签内，并附在被标示的产品上，以被高层的信息处理软件识别、传递和查询，进而在互联网的基础上形成专为供应链企业服务的各种信息服务。

2. PML 信息服务器

PML(Physical Markup Language)信息服务器由产品生产商建立并维护，储存着这个生产商生产的所有商品的文件信息。根据事先规定的原则对产品进行编码，并利用标准的 PML 对产品的名称、生产厂家、生产日期、重量、体积、性能等详细信息进行描述，从而生成 PML 文件。在 PML 文件中，除了包括不会改变的产品信息(如物质成分)外，还包括经常性变动的数据(动态数据，如输送过程中的水果温度)和随时间变动的数据(时序数据，如物品所处的地点)。每件产品的 PML 文件都存储在一个 PML 信息服务器中，是一个专用的计算机，为其他计算机提供需要的文件。可见，PML 信息服务器在物联网中的作用在于以通用的格式提供对产品原始信息的描述，便于其他结点访问。

PML 是一种用于描述物理对象、过程和环境的通用语言，其主要目的是提供通用的标准化词汇表，来描绘物体的相关信息。PML 以可扩展标记语言 XML 的语法为基础，正被 Auto-ID 中心开发成一种开放的标准，这样全世界任何地方的供应商就可以以一种能被理解的统一高效的方式来描述产品的信息。PML 核心提供通用的标准词汇表来分配直接由 Auto-ID 的基础结构获得的信息，如位置、组成以及其他遥感勘测的信息。

PML 分为两个主要部分：PML 核心与 PML 扩展。PML 核心提供通用的标准词汇来分配直接由 Auto-ID 基础结构获得的信息，如位置、组成及其他遥感勘测的信息。PML 扩展用于将非 Auto-ID 基础结构产生的或其他来源合成的信息结合成一个整体。

第一个实现的扩展是 PML 商业扩展，包括丰富的符号设计和程序标准，使组织内或组织间的交易得以实现。

PML 不但提供了描述自然物体、过程和环境的标准，并可供工业和商业中的软件开发、数据存储和分析工具之用。PML 也提供了一种动态的环境，使与物体相关的、静态的、暂时的、动态的和统计加工过的数据互相交换。因为 PML 将会成为描述所有自然物体、过程和环境的统一标准，所以，它的应用会非常广泛，并且会进入到所有行业。当所有商品的信息都可通过 PML 文件得到时，公司就能够以革新方式利用这些数据。例如，公司可以在文件中设置一个触发器，以便有效期将要结束时降低产品价格；第三方物流提供者则有可能在合同中表示可随时为客户提供货物在运输过程中的温度，使客户可以监督货物运输的质量。

一个典型的 PML 服务器的原理图如图 11.2 所示，图中各模块的功能描述如下：

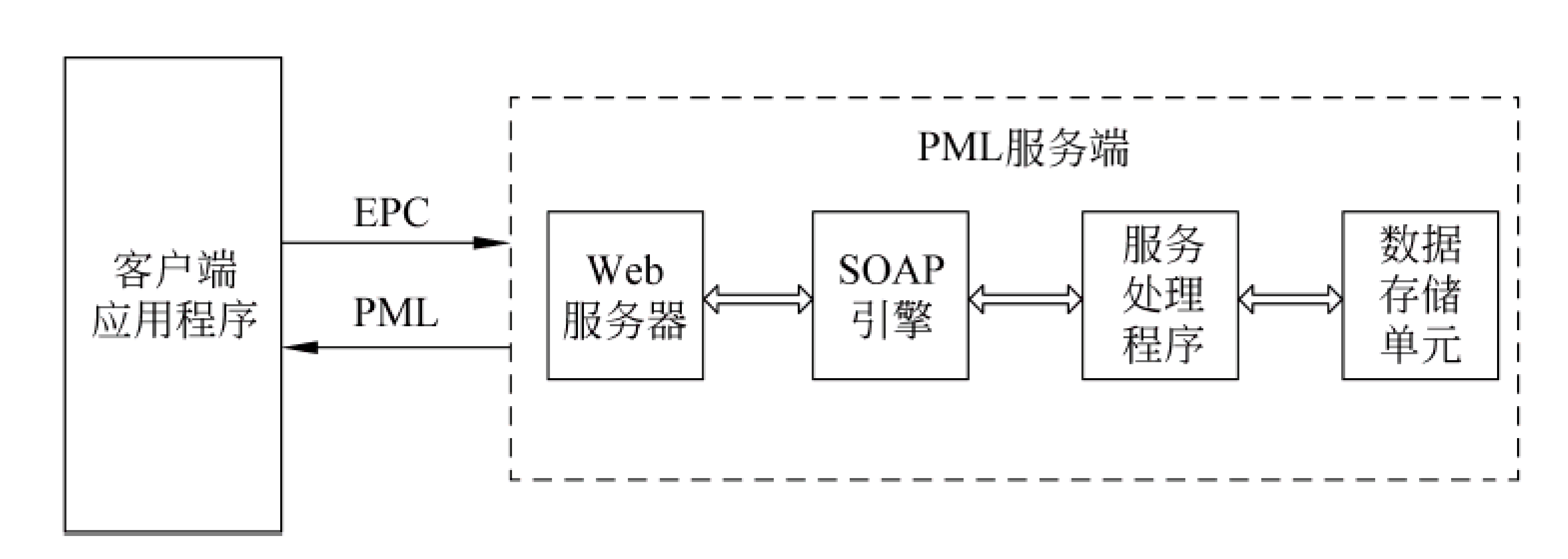


图 11.2 PML 服务器原理图

(1) Web 服务器。它是 PML 信息服务中唯一直接与客户端交互的模块，位于整个 PML 信息服务的最前端，可以接收客户端的请求，进行解析、验证，确认无误后发送给 SOAP 引擎，并将结果返回给客户端。

(2) SOAP 引擎。它是 PML 信息服务器上所有已部署服务的注册中心，可以对所有已部署的服务进行注册，提供相应组件的注册信息，将来自 Web 服务器的请求定位到相应的服务器处理程序，并将处理结果返回给 Web 服务器。

(3) 服务器处理程序。它是客户端请求服务的实现程序，包括实时路径更新程序、路径查询程序和原始信息查询程序等。实时路径更新程序用来更新产品在供应链的路径信息，包括单位角色、单位名称、仓库号、读写器号、时间、城市等；路径查询程序用来查询产品在供应链中的路径信息；原始信息查询程序用来查询产品的原始信息，该信息通常以 PML 文件形式存储在数据存储单元中。产品每一个服务处理程序完成一项客户端提出的具体请求，它接受客户端送过来的参数，完成一些逻辑处理和数据存取操作，并将处理结果返回给 SOAP 引擎。

(4) 数据存储单元。它用于 PML 信息服务器端数据的存储，主要用于客户端请求数据的存储，存储介质包括各种关系数据库或者一些中间文件，如 PML 文件。存取的数据包括产品级数据和个体级数据两类：产品级数据记录产品的规格、性能、几何特性等，这些数据在产品中是公有的信息；个体级数据记录个体在供应链中流动时所特有的历史信息(地点、时间戳、传感测量值)，以及产品的个性化参数等。

在整个物联网系统中，PML 信息服务主要提供以下 3 个功能：

(1) 实时路径信息的存储。主要用于当产品经过供应链成员结点,被读写器捕获时,将此时的状态信息收集,并通过产品 EPC 立刻传入与产品对应的 PML 信息服务器上,以供定位跟踪或其他用途时查询。

(2) 产品路径信息查询。实现产品从生产商、分销商、批发商、零售商到最终用户等供应链各成员结点的路径信息跟踪显示。通过电子标签实现对产品的实时跟踪、产品物流控制和管理,各成员可以根据产品路径来推测产品的来源渠道,并判别产品真伪,同时也可以据此灵活调节自己的库存,大大提高供应链的运行绩效。

(3) 产品原始信息查询。主要用于查询产品 EPC 对应的产品出厂时的原始信息,这项信息可以和路径信息结合作为产品防伪的一项重要措施。

3. ONS

ONS 的作用是在各信息采集结点与 PML 信息服务器之间建立联系,实现从 EPC 到 PML 信息之间的映射。读写器识别 RFID 标签中的 EPC 编码,ONS 则为带有射频标签的物理对象定位网络服务,这些网络服务是一种基于 Internet 或者 VPN 专线的远程服务,可以提供和存储指定对象的相关信息。实体对象的网络服务通过该实体对象的 EPC 代码进行识别,ONS 帮助读写器或读写器信息处理软件定位这些服务。

ONS 定位的网络服务可以将 EPC 关联到与物品相关的 Web 站点或其他 Internet 资源。当前典型的 ONS 服务用来定位某一 EPC 对应的 PML 信息服务器。ONS 服务是联系前台 Savant 软件和后台 PML 信息服务器的网络枢纽,并且 ONS 的设计与架构都以域名解析服务 DNS 为基础。因此,可以使整个 EPC 网络以 Internet 为依托,迅速架构并顺利延伸到世界各地。ONS 的体系结构如图 11.3 所示,是一个分布式的系统架构,主要由以下几部分组成。

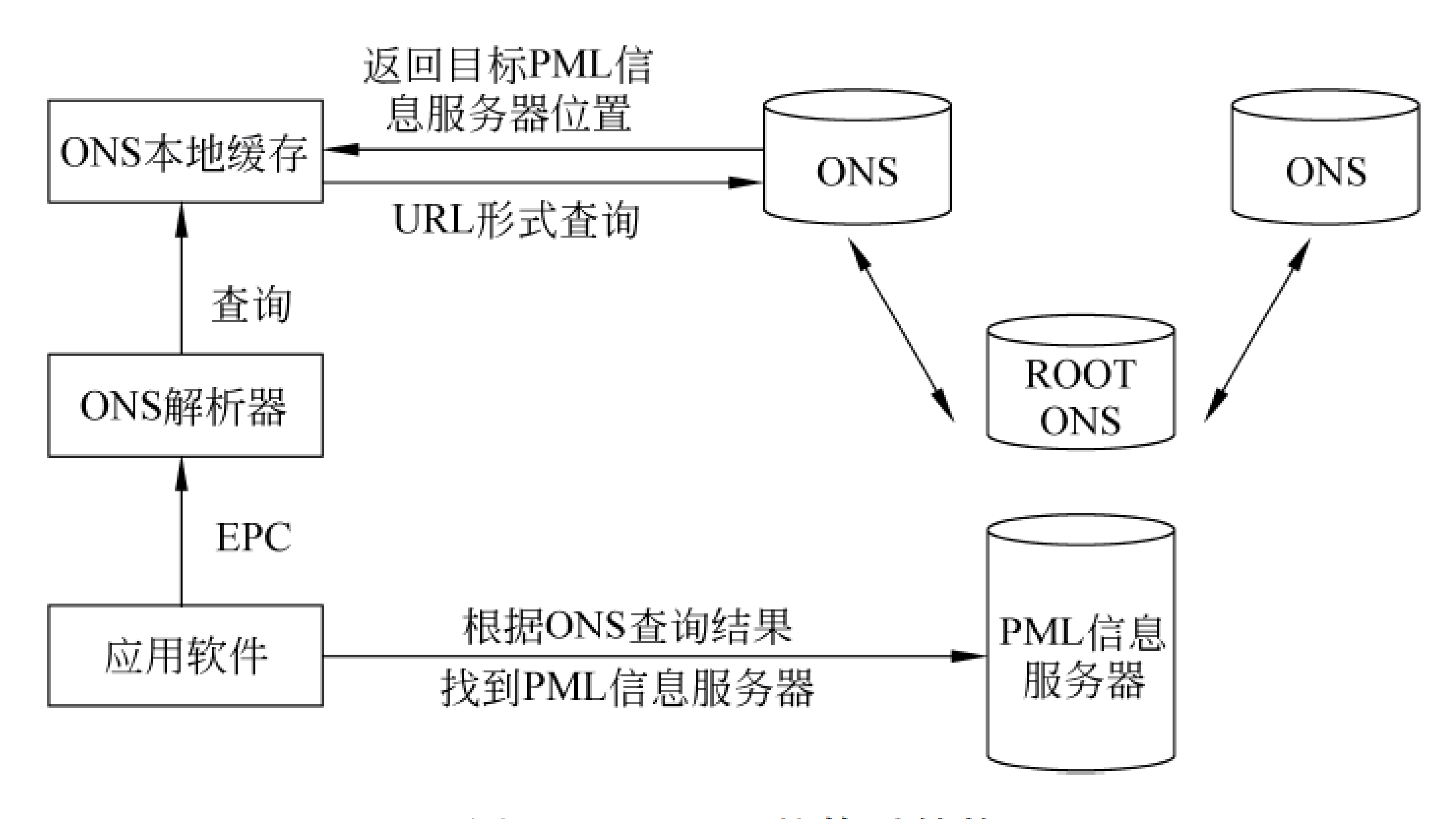


图 11.3　ONS 的体系结构

(1) 映射信息。映射信息以记录的形式表达了 EPC 编码和 PML 信息服务器之间的一种映射,它分布式地存储在不同层次的 ONS 服务器里。

(2) ONS 服务器。如果某个请求要求查询一个 EPC 对应的 PML 信息服务器的 IP 地址,则 ONS 服务器可以对此作出响应。每一台 ONS 服务器拥有一些 EPC 的授权映射信息和 EPC 的缓冲存储映射信息。

(3) ONS 解析器。ONS 解析器负责 ONS 查询前的编码和查询语句格式化工作，它将需要查询的 EPC 转换为 EPC 域前缀名，再将 EPC 域前缀名与 EPC 域后缀名结合成一个完整的 EPC 域名，最后由 ONS 解析器发出对这个完整的 EPC 域名进行 ONS 查询的请求，获得 PML 信息服务器的网络定位。

(4) ONS 本地缓存。ONS 本地缓存可以将经常查询和最近查询的"查询-应答"值保存在内，作为 ONS 查询的第一入口点，这样可以减少对外查询的数量，提高本地相应效率，减少 ONS 服务器的查询压力。ONS 本地缓存同时也用于响应企业内部 ONS 的查询，用于物品跟踪。

4. Savant 系统

Savant 系统在物联网中处于读写器和企业应用程序之间，相当于物联网的神经系统。Savant 系统采用分布式结构，层次化组织、管理数据流，具有数据搜集、过滤、整合与传递等功能，因此，能将有用的信息传送到企业后端的应用系统或者其他 Savant 系统中。

各个 Savant 系统分布在供应链的各个层次结点上，如生产车间、仓库、配送中心及零售店，甚至在运输工具上。每一个层次上的 Savant 系统都将收集、存储和处理信息，并与其他 Savant 系统进行交流。例如，一个运行在商店的 Savant 系统可能要通知分销中心还需要其他的产品，在分销中心的 Savant 系统则通知一批货物已经于一个具体的时间出货了。

可以看到，EPCglobal 网络的架构其实是一个应用层的架构。

11.2 EPCglobal 网络安全

11.2.1 EPCglobal 网络的安全性讨论

关于 EPCglobal 网络本身的安全性的研究目前还不多见，多数文献还是重点讨论在超高频第二代 RFID 标签上如何实现双向认证协议。

文献[5]认为 ONS 架构存在严重的安全缺陷，提出了一种基于分布式 Hash 表 DHT 的 P2P 名字服务架构，并在 PlantLab 平台上进行了研究，文中称这个架构为 OIDA。与 ONS 架构相比，该架构提供了多方面的安全，并提高了性能和可扩展性。

文献[6]设计了一种 RFID 标签和读写器之间的双向认证协议，且该协议可以在 EPCglobal 兼容的标签上使用。该协议可以提供前向安全(forward security)。该文还提出了一种 P2P 发现服务的 EPC 数据访问方法，该方法比基于中央数据库的方法具有更好的可扩展性。

文献[7]指出在 EPCglobal 的 1 类 2 代(Class-1 Generation-2)RFID 标签中，标签的标识是以明文进行传输的，于是很容易被追踪和克隆。通过对称和非对称密码加密的方法在廉价标签中可能不太可用。虽然一些针对第 2 代标签的轻量级的认证协议已经提出，但这些协议的消息流与第 2 代标签的消息流不同，因而存在的读写器可能不能读新的标签。该文提出一种新的认证协议，称为 $Gen2^+$，该协议依照第 2 代标签的消息流，提供了向后兼容性。该协议使用了共享的假名(pseudonym)和循环冗余校验(CRC)来获得读

写器对标签的认证。并利用读存储命令来获得标签对读写器的认证。论证结果表明 Gen2$^+$ 在跟踪和克隆攻击下更加安全。

文献[8]分析了第 2 代标签的安全缺陷,包括泄漏、对完整性的破坏、拒绝服务攻击以及克隆攻击。泄漏威胁即不该读取标签的读写器读了标签信息。广义来说,泄漏威胁指在 RFID 标签信息在保存在标签和在传递给读写器的过程中被泄漏。拒绝服务攻击就是当标签被访问时被敌对读写器阻塞了,即当一个读写器需要读标签信息时,被另一个敌对读写器阻止了这种访问。这种阻塞可能是持续的,导致标签信息总是无法被读取。破坏完整性威胁是指非授权地对标签存储的信息或者传递给读写器的信息进行修改。克隆攻击就是某个非法标签的敌对行为导致欺骗读写器,使读写器以为自己正在与某个设备进行正确的信息交换。这种攻击中,仿真程序或硬件在一个克隆标签上运行,伪造了读写器期望的正常的操作流程。为了应对这些威胁,可采用的方法包括使用会话来避免泄漏,引入高密度(dense)读写器条件来避免拒绝服务攻击,组合安全协议和空中接口、Ghost Read 以及 Cover coding 方法来克服破坏完整性攻击。对于克隆攻击还没有好的防御方法。

文献[9]提出了在基于 RFID 的供应链中可以检索和分析分布式的 EPC 事件数据,组合这些数据可能导致商业机密的泄漏。该文提出了一个基于证书(license)的访问控制原型系统,来保护商业方的隐私。该方法依照了欧盟提出的隐私保护设计(privacy-by-design)原则,可以减少暴露的数据。

11.2.2 EPCglobal 网络中的数据清洗

由于读写器异常或者标签之间的相互干扰,有时采集到的 EPC 数据可能是不完整的或错误的,甚至出现多读和漏读的情况。漏读(false negative)是指:当一个标签在一个阅读器阅读范围之内时,该阅读器没有读到该标签。多读(false positive)是指:当一个标签不在一个阅读器阅读范围之内时,该阅读器仍然读到该标签。如果将源数据直接投入到实际应用中,得到的结果一般都没有应用价值,所以在对 RFID 源数据进行处理之前,需要对数据进行清洗。Savant 要对读写器读取到的 EPC 数据进行处理,消除冗余数据,过滤掉“无用”信息,以便传送给应用程序或上级 Savant 的是“有用”信息。下面先简要介绍一下“多读”的冗余数据清洗方法。

冗余数据的产生主要是由于以下两个因素:第一,在短期内同一台读写器对同一个数据进行重复上报。如在仓储管理中,对固定不动的货物重复上报;在进货、出货过程中,重复检测到相同物品。第二,多台临近的读写器对相同数据都进行上报。读写器存在一定的漏检率,这和读写器天线的摆放位置、物品离读写器远近、物品的质地都有关系。通常为了保证读取率,可能会在同一个地方相邻地摆放多台读写器,这样,多台读写器将监测到的物品上报时,就可能会出现重复。

另外,很多情况下用户可能还希望得到某些特定货物的信息、新出现的货物信息、消失的货物信息或只是某些地方的读写器读到的货物信息。用户在使用数据时,希望最小化冗余,尽量得到靠近需求的准确数据。

冗余信息的解决办法是设置各种过滤器进行处理。可用的过滤器有很多种。典型的

过滤器有 4 种：产品过滤器、时间过滤器、EPC 过滤器和平滑过滤器。产品过滤器只发送与某一产品或制造商相关的产品信息，也就是说，过滤器只发送某一范围或方式的 EPC 数据；时间过滤器可以根据时间记录来过滤事件，如一个时间过滤可能只发送最近 10min 内的事件；EPC 过滤器可过滤符合某个规则 EPC；平滑过滤器可处理出错的情况，包括漏读和错读。

对于"漏读"的情况，需要通过标识之间的关联度(如同时被读到)来找回漏掉的标识。文献[10]提出基于对监控对象动态聚簇概念的 RFID 数据清洗策略，通过有效的聚簇建模和高效的关联度维护来估算真实的小组，这里所谓"小组"就是常常会同时读取的具有某种关联度的标签。在估算真实的小组基础上进行有效的清洗。由于引入了新的维度，在有小组参与的情况下，无论数据量的大小和组变化的程度，与考虑时间维的相关工作相比，该文提出的模型可以有效地利用组间成员的关系来提高清洗的准确性。

研究与思考

[1] EPCglobal 网络中的供应链 RFID 协议，与基于连线数据库后端的 RFID 协议相比具有不同的网络结构，这会对安全协议的设计导致什么不同。

[2] EPCglobal 中的开放的 ONS 架构是否安全，是否可能存在一种类似于 DNS 攻击的 ONS 攻击。

[3] EPCglobal 网络中是否存在隐私泄漏问题，如何设计保护隐私的方案。

进一步阅读建议

[1] EPCglobal 中国，http://www.epcglobal.org.cn/.

本章参考文献

[1] Auto-ID Center，AutoID 实验室[OL]，http://www.autoidlabs.org.

[2] EPCGlobal[OL]，http://www.epcglobalinc.org.

[3] EPCglobal，EPC Radio-Frequence Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860MHz ～ 960MHz，Version 1.1.0[S]，Dec. 2005，http://www.epcglobalinc.org/standards/uhfc1g2/uhfc1g2_1_1_0-standard-20071017.pdf.

[4] 杨刚，沈沛意，郑春红等. 物联网理论与技术[M]. 北京:科学出版社，2010.9.

[5] B. Fabian，Implementing Secure P2P-ONS[C]，In Proc. of IEEE International Conference on Communications (ICC09)，pp. 1-5，14-18 June 2009.

[6] Tieyan Li，R. Deng，Scalable RFID authentication and discovery in EPCglobal network[C]，In Proc. of Third International Conference on Communications and Networking in China(ChinaCom 2008)，pp. 1138-1142，25-27 Aug. 2008.

[7] Hung-Min Sun；Wei-Chih Ting，A Gen2-Based RFID Authentication Protocol for Security and Privacy[J]，IEEE Transactions on Mobile Computing，vol. 8，no. 8，pp. 1052-1062，Aug. 2009.

[8] A. Razaq，Wai Tong Luk，Kam Man Shum，Lee Ming Cheng，Kai Ning Yung，Second-Generation

RFID[J], IEEE Security & Privacy, vol. 6, no. 4, pp. 21-27, July-Aug. 2008.

[9] Matthieu-P. Schapranow, Martin Lorenz, Alexander Zeier, Hasso Plattner, License-based Access Control in EPCglobal Networks[C], In Proc. of 7th European Workshop on Smart Objects: Systems, Technologies and Applications (RFID SysTech 11), pp. 1-7, 17-18 May 2011.

[10] 谷峪,于戈,胡小龙,王义. 基于监控对象动态聚簇的高效 RFID 数据清洗模型[J]. 软件学报, 2010. 4.

第 12 章　基于无线体域网的远程医疗安全

物联网的一个基本应用就是提供无处不在的远程医疗服务或者是可穿戴的医疗计算服务。远程医疗(Telehealth)也是(除 M2M 技术外)美国最受关注的两个物联网应用之一(另一个是智能电网)。这其中的一个典型形式是通过无线体域网的传感器结点采集体表(或者体内)的健康监控数据,借助网关(如智能手机)将这些数据发送到医疗机构,以得到健康状况的监控,并实现远程医疗诊断和指导。

一个实际的例子就是 2010 年被 readwriteweb 网站评为物联网 10 大进展之一的 Nike+鞋。美国耐克公司和苹果公司合作推出 Nike iPod 系列产品,首次将运行与音乐完美结合起来。这两家公司全作开发的首款产品为 Nike iPod 运动组件,包括一个内置于鞋中的传感器和一个与 iPod 连接的接收器。通过 Nike iPod 运动组件,就可以将 Nike 运动鞋与 iPod nano 连接,iPod 就可以存储并显示运动的时间、距离、热量消耗值和步幅等数据。使用都也可以通过耳机了解这些实时数据。跑步后将 iPod 与电脑相连开始同步,就可以完美地同步你的跑步和健身数据。通过直观漂亮的图形界面查看每次、每周或者每月的跑步速度、距离或者消耗的卡路里,每次健身的数据都一目了然。还可以分析你的成绩,看看是否有打破上一次的记录,或者看看距离自己设定的目标还有多远。该产品甚至还有自己的社交网络:它可以自动发布状态消息。

另外,据 2012 年初的媒体报道,苹果公司正在研发能戴在用户手腕上的计算机,谷歌公司正在开发一款运行 Android 的眼镜产品,配置的小型显示屏位于眼睛前方数英寸处,具有 3G 或 4G 连接功能,包括动作和 GPS 在内的大量传感器。因此,我们认为,可穿戴的远程医疗计算可能在不久的将来广泛商用。

12.1　无线体域网概述

无线体域网(Wireless Body Area Networks,WBAN)是由大量安置于人体体表或体内的传感器结点通过自组织方式组成的无线传感器网络(Wireless Sensor Networks,WSN),用于检测人体生理数据或周边状况信息。无线体域网在日常生活、医疗、娱乐、军事、航空等领域有重要的应用。为了适应无线体域网的应用需求,2007 年 IEEE 成立 802.15.6[1]工作小组,旨在制订用于无线体域网的通信标准。无线体域网还有其他的称呼,如体域传感网(Wireless Body Sensor Network,WBSN 或 BSN)、生物医疗传感器网络(Biomedical Sensor Network,BSN)和无线体域传感网(Wireless Body Area Sensor Network,WBASN)。

WBAN 综合了 WPAN(无线个域网)、WSN(无线传感器网)、无线短距离通信等各种技术。一般认为 WBAN 是 WPAN 的一种延伸。一般意义上的 WPAN 是指将诸如 PC、

智能手机、平板电脑、MP3、数码相机等个人通信设备组成小型的网络。而 WBAN 将这些设备扩展到了放置在人体上和(或)人体内的各类传感器。另外,WBAN 可以看做是 WSN 在远程医疗诊断和监护中的一个非常重要的应用(它同时也是物联网的一个重要应用)。短距离无线通信技术是短距离传感器和终端设备之间及终端和终端之间通信的主要方式。和其他短距离无线通信技术相比,WBAN 需要的短距离通信技术在相同的功率下,数据传输速率更高;或者在相同的数据传输速率下,需要的功率更低。图 12.1 给出一个 WBAN 的典型应用场景。

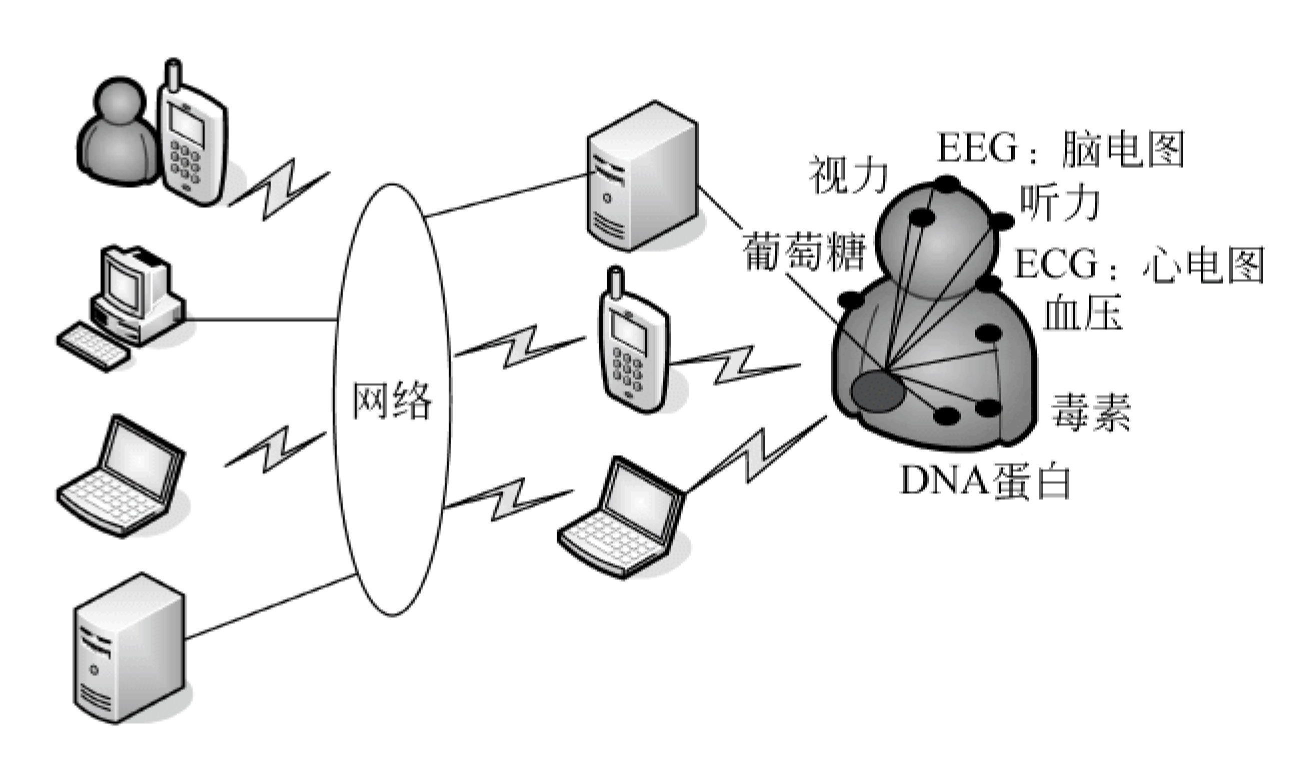


图 12.1 WBAN 的典型应用场景

12.1.1 无线体域网的系统架构

目前多采用先分布式采集或感知、再集中式处理的方式。考虑到网络规模较小,并且每对传感器结点之间的通信也不是必需的,因此分布式采集部分通常采用星形拓扑结构。为了更好地概括和总结各种已有的 WBAN 的系统架构,通过兼容性地整合这些结构给出一种全面的架构形式,如图 12.2 所示,WBAN 的系统架构包含 3 个层次(在文献[3]基础上进行了修改):

(1) 第 1 层包含一组具有检测功能的传感器结点或设备。在医疗领域,传感器能够测量和处理人体的生理信号或所在环境的信息,然后将这些信息传送给外部控制结点或头结点(Head),还可以接收外部命令以触发相应动作。在非医疗领域,可穿戴的设备(如耳机、MP3 播放器和游戏控制器等)都可以包含进来。

根据相对于人体所在位置可将传感器结点分为 3 类:

① 可植入体内的传感器结点,包括可植入的生物传感器和可吞入的传感器(如摄像药丸)。

② 可穿戴在身体上的传感器结点,如血压传感器、血氧饱和度传感器和温度传感器等。

③ 在身体周围并距离身体很近(或较近)的用于识别人体活动或行为的周围环境结点。

(2) 第 2 层是 WBAN 头结点或主结点(Master Nodes),可能还包括汇聚(Sink)结点、基站(Base Station)、网关。它负责和外部网络进行通信,并临时存储从第 1 层收集上

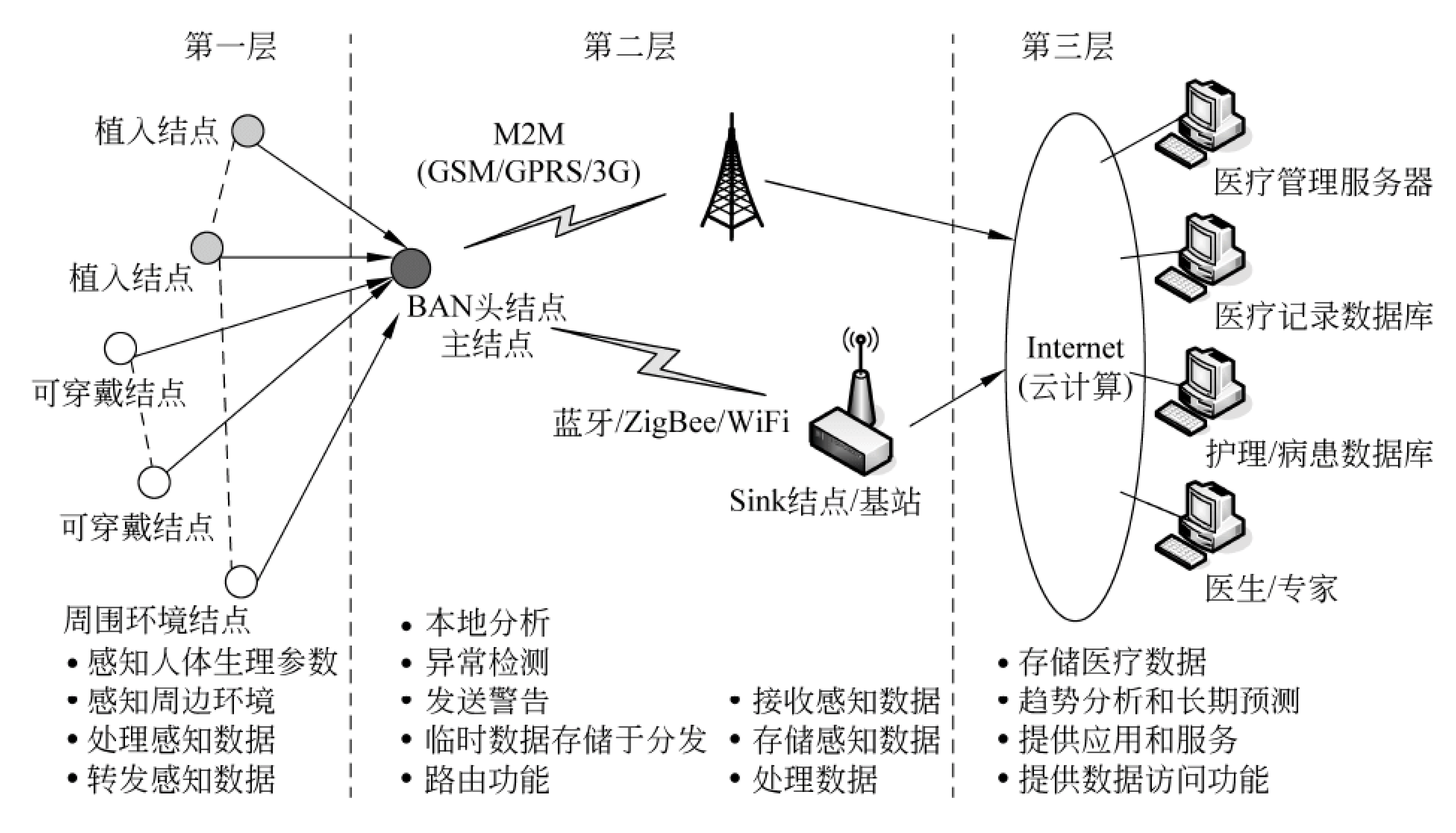


图 12.2　WBAN 系统架构

来的数据。它以低能耗的方式管理各个传感器结点或设备，接收和分析感知数据及执行规定的用户程序。基站可以是资源相对丰富的智能手机，或能够上网的手持设备。一般情况下，当 WBAN 网络结点个数不多时，H 结点（即头结点）可与汇聚（Sink）结点合二为一，只有在 WBAN 网络结点较多的情况下会存在多个头结点（即进一步分层的头结点，也称为簇头结点，Cluster Head），此时就需要一个 Sink 结点或基站来负责收集由这些头结点发送的信息，并作为网关与外部网络进行连接。

（3）第 3 层是包括提供各种应用服务的远程服务器的外部网络，例如，医疗服务器保留注册用户的电子医疗记录，并向这些用户、医务人员和护理人员提供相应服务。我们认为，第 3 层和云计算结合起来，是比较有前途的，它可以满足大量 WBAN 感知数据的保存问题，便于多用户多地点访问保存的数据。

12.1.2　无线体域网的特征

1. 与无线传感器网络的区别

虽然无线体域网 WBAN 与无线传感器网络 WSN 具有一定相似性，如分层汇聚网络架构，传感器结点等，但两者仍有区别，表现在如下方面（这些区别也构成了安全设计的约束条件）。

（1）WBAN 是由不多于几十个传感器结点组成的无线移动网络，网络的规模较小。

（2）网络内部采集业务（生理）数据的不同，该网络具有业务数据多样性的特点。

（3）WSN 中大多数传感器结点是固定不动的，而 WBAN 需要考虑人体对整个网络的影响，人体作为整个网络传感器结点的载体，人体四肢的运动会导致网络拓扑结构的变化。

（4）与 WSN 相比，WBAN 对能量的要求更加苛刻，其发射功率要求更低。WBAN 网络协议设计的首要目标是保证能量的高效利用。

2. 典型的医疗传感器结点

对无线体域网中终端结点的了解是安全设计时必须考虑的因素。无线体域网的一个典型应用是医疗领域。随着人们对各种健康保健和健康预测的要求越来越高,各种医疗传感器不断涌现。这些传感器被广泛用于外科手术设备、医疗保健和家庭护理中。

根据传感器在身体的位置不同,可将医疗传感器分为以下三大类:植入传感器,与体液接触的可穿戴传感器和无接触可穿戴传感器。下面简要介绍这三大类传感器。

(1) 植入传感器。植入传感器体积小、重量轻,同时其功率必须非常小。更重要的是,它们不能随着时间的推移而衰变。对功率的要求是植入传感器正常工作所面临的主要挑战之一。例如,压电聚合传感器体积小,可靠性高,不需要外部动力而且能长时间持续工作。这类传感器可应用于监视病人活动的心脏起搏器。

(2) 与体液相接触的可穿戴传感器。如一次性血压传感器,可用外科手术和重症监护,以便持续地监控病人的血压情况。这是在给病人进行静脉输液的同时测量其血压的最理想方式。这类传感器需要每 24 个小时更换一次,以保证传感器的清洁卫生。这类传感器被连到一个监控器上,以便记录下所有的信息。

(3) 无接触可穿戴传感器。在体液不能和传感器相接触的情况下,可将无接触可穿戴式传感器应用于医疗设备。大多数情况下,这类传感器都不是一次性的,医院或是家庭护理均可使用这类传感器。如用于研究睡眠窒息的压电(或热电)传感器;用于测量体表/体内温度的温度传感器。

3. 无线体域网的特点

(1) 规模小,可扩展,近距离,以人体为中心的网络。

由于人体几何结构的限制,WBAN 的网络规模很小,而且传感器结点根据所采集的数据信息放置在身体的相应部位,其相对位置固定。如果数据采集点的数量较少,则 WBAN 的规模可能很小。随着人们对各种服务的需求增多,在 WBAN 中也会随时增加多种微型的便携式设备(新的数据采集点),因而 WBAN 应该是可扩展的。另外,由于人体本身的几何特性,导致传感器结点之间的通信范围是有限的,WBAN 的通信距离通常是 2～5 米。因而针对这样近距离的网络来说,用于无线传感器网络,无线自组织网络等无线网络中的一些网络协议和算法(如基于分簇拓扑结构、多层结构等)可能就不再适用于无线体域网。无线体域网的网络拓扑结构和路由算法相对比较简单,普通传感器结点最多通过三跳(Hop)路由就可以将采集的数据传送到汇聚结点,大多数情况下是单跳星型网络。

与其他传统网络最大的不同是:无线体域网是以人体为中心的网络,还必须考虑通讯对人体健康的安全性,这就要求传感器结点的发射功率必须足够低(这也是当今“绿色通信,Green Communication”的一个基本要求);另外,各传感器结点或便携式设备以人为载体,分布在人体的不同位置,人体不仅可以影响传感器结点之间的通信,而且也可以作为传感器结点之间的通信信道。这就使得无线体域网的信道特征变得复杂,在网络结构设计中必须考虑这一点。这些特点也是安全设计时需要考虑到的。

(2) 数据业务多样性、数据相关性的网络。

无线体域网的数据业务多样性包括以下两方面的含义:一方面是该网络可以提供多

种业务服务，如数据业务、音频、视频、Internet 服务等，这些业务可以同时由网络提供以满足用户的需求；另一方面，对于同一种应用场景，如医疗应用，无线体域网中所处理的数据会因为传感器结点所采集的人体生理数据的不同（如体温、心跳、血压等）而呈现数据多样性的特点。

在无线体域网的医疗应用中，采集的各种生理信息在一定程度上还具有相关性。比如，当一个病人发烧时，那么其体温、血压、心率、呼吸率等生理参数会相应升高，这种生理参数的相关性会直接影响到整个网络传输业务的相关性。因此，在传输数据时必须保证多种生理数据的传输在时间上能够同步，以使得传输到远程中心的数据是同一时间域采集的数据，便于准确地掌握病人的身体状况。另外，生理数据一般在某个区间范围内变动，这便于数据的压缩传输（并在数据保密增强时可以加以利用）。

（3）动态拓扑结构的网络。

无线体域网的拓扑结构可能因为下列因素而改变：身体四肢的随机移动会导致结点之间的通信中断；电能耗尽造成结点出现故障或失效；新结点的加入。这要求无线体域网必须能够适应这些变化，因而网络拓扑是可重构的。

（4）以数据为中心（Data Centric）的网络。

和无线传感器网络相同，无线体域网也是一个以数据为中心的网络。无线体域网中的结点采用结点编号标识，远程控制中心如果需要某方面的数据信息时，只需要把指令直接告诉网络的汇聚结点，然后由汇聚结点向全网广播该指令，并不需要给特定编号的结点发送请求信息。

12.2　WBAN 安全分析

12.2.1　WBAN 的安全威胁

作为无线网络的一种，常见的无线网络中的安全威胁如窃听、篡改、伪造数据注入、拒绝服务攻击等都依然存在。针对 WBAN 的特点，其中比较重要的安全威胁包括以下两方面。

（1）来自结点妥协的威胁：WBAN 中的结点容易被俘获，所以结点中的加密数据和密钥不能在一起存放，否则一旦结点被捕获，数据就会暴露。

（2）来自网络变化的威胁：WBAN 的网络是高度变化的，结点可能加入或离开网络，结点可能因缺电而失效，攻击方甚至可以伪造虚假结点。

由于 WBAN 在医疗应用方面的特殊性，主要的安全需求归纳如下。

（1）保密性：为阻止患者的医疗数据不被泄漏，数据需要保密。这种保密方法需要能够对抗结点的妥协。例如容入侵的加密，抗密钥泄漏的加密（Resilient to Key Leakage[4]）等。

（2）私有信息保护：个人的医疗信息的私有性是受法律保护的（美国颁布了 Health Insurance Portability and Accountability Act，即 HIPAA 条例[5]），因此对医疗信息的检索应该制定严格的访问控制策略，信息检索应该是能够保护私有信息的信息检索

(Privacy-aware Information Retrieval)[6]。

(3) 完整性：防止数据的篡改是至关重要的，这需要保证数据的完整性，而且完整性安全方案需要考虑到数据的动态更新。

此外，网络的可靠性、健壮性、可用性对于医疗应用来说也是十分重要的。无线传感器网络安全的角度总结了 WBAN 的安全威胁与防御方法，如表 12.1 和表 12.2 所示[7]。

表 12.1 WBAN 从 OSI 层次分析的威胁、DoS 攻击与防御

协议层	DoS 攻击	防 御 方 法
物理层	阻塞(Jamming)	扩展频谱，优先消息，更短的责任周期，区域映射，模式改变
	干扰(Tampering)	抗干扰，隐藏
链接层	(包)冲突	纠错码
	非公平竞争(信道)	使用较小的帧
	耗尽(信道资源)	流量限制
网络层	怠慢和贪婪	冗余，探测
	汇聚结点攻击	加密
	方向误导	输出流量的过滤，认证和监控
	黑洞	认证，监控，冗余
传输层	泛洪	客户端谜题
	破坏同步	认证

表 12.2 WBAN 的安全威胁和解决方法

安 全 威 胁	安 全 需 求	可能的安全方案
未认证和非授权的访问	密钥建立和信任建立	随机密钥分发，公钥密码学
消息泄漏	机密性和隐私性	链接层和网络层加密，访问控制
消息篡改	完整性和认证性	有密钥的安全 Hash 函数，数字签名
拒绝服务攻击	可用性	入侵检测，冗余
结点捕获和结点妥协	抗结点妥协(compromise)	检测不一致性，结点回收，抗干扰
路由攻击	安全路由	安全路由协议
入侵和高层安全攻击	安全组管理，入侵检测，安全数据聚集	安全组通信，入侵检测

12.2.2 WBAN 的安全方案简介

目前对 WBAN 安全方面的研究尚处于起步阶段，这里只简要介绍少量研究成果。在加密算法方面，Y. Law 等认为能量效率最高的加密方法是 Rijndael，他们测试了在密钥建立和加密、解密操作的 CPU 周期数，比较的结果如表 12.3 所示[8]。

表 12.3　不同加密方法的性能比较

密钥建立						
排序	大小优化			速度优化		
	代码空间	数据空间	速度	代码空间	数据空间	速度
1	RC5-32	MISTY1	MISTY1	RC6-32	MISTY1	MISTY1
2	KASUMI	Rijndael	Rijndael	KASUMI	Rijndael	Rijndael
3	RC6-32	KASUMI	KASUMI	RC5-32	KASUMI	KASUMI
4	MISTY1	RC6-32	Camellia	MISTY1	RC6-32	Camellia
5	Rijndael	RC5-32	RC5-32	Rijndael	Camellia	RC5-32
6	Camellia	Camellia	RC6-32	Camellia	RC5-32	RC6-32
加密(CBC/CFB/OFB/CTR)						
排序	大小优化			速度优化		
	代码空间	数据空间	速度	代码空间	数据空间	速度
1	RC5-32	RC5-32	Rijndael	RC6-32	RC5-32	Rijndael
2	RC6-32	MISTY1	MISTY1	RC5-32	MISTY1	Camellia
3	MISTY1	KASUMI	KASUMI	MISTY1	KASUMI	MISTY1
4	KASUMI	RC6-32	Camellia	KASUM	RC6-32	RC5-32
5	Rijndael	Rijndael	RC5-32	Rijndael	Rijndael	KASUMI
6	Camellia	Camellia	RC6-32	Camellia	Camellia	RC6-32

利用生理数据产生随机密钥进行认证或者保密通信是 WBAN 安全研究的主流方向，目前有一些论文探讨了这个问题。Poon 等[9]建议使用脉搏间隔(Inter-Pulse Intervals，IPIs)作为一个生物特征进行身份识别或者对称密钥分发。S. D. Bao 等[10]建议使用心跳信息进行身份识别。K. Venkatasubramanian[11]提出基于血流图(Plethysmogram)的信息处理产生对称密钥生成方法 PKA(PPG based Key Agreement)，以及通过心电图(Electrocardiogram，ECG)信号进行密钥协商的方案 EKA(Ekg-based Key Agreement)[12]。M. R. Kanjee 等[13]提出基于生理学的认证方案。

研究与思考

[1] WBAN 作为物联网远程医疗应用的一个局域子网，可利用 M2M 技术作为网关，提供远程连接和远程访问功能。结合 Android 手机应用开发，能否开发一个 WBAN 应用。

[2] WBAN 和可穿戴计算(wearable computing)的结合可以导致很多新颖的应用。设想一下还可以有什么样的创新应用，如嵌入传感器到帽子、衣服、眼镜之类。

[3] WBAN 与云计算的结合，可以发挥端和云的各自优势，具有良好的应用前景。这种 WBAN-Cloud 架构下的安全性应该注意什么。

进一步阅读建议

WBSN 的安全机制的设计往往与具体应用场景或者人体参数高度相关，借助应用场景或者人体提供的某些安全参数进行认证或者密钥生成是一个思路。

[1] S. D. Bao，Y. T. Zhang，and L. F. Shen. A Design Proposal of Security Architecture for Medical Body Sensor Networks[C]，In Proc. of the 3rd International Workshop on Wearable and Implantable Body Sensor Networks (BSN06)，Cambridge，Massachusetts，USA，2006，84-87.

[2] F. M. Bui and D. Hatzinakos，Biometric Methods for Secure Communications in Body Sensor Networks：Resource-efficient Key Management and Signal-Level Data Scrambling[J]，EURASIP Journal on Advances in Signal Processing，2008.

[3] N. Challa，H. Cam，and M. Sikri，Secure and Efficient Data Transmission over Body Sensor and Wireless Networks [J]，EURASIP Journal on Wireless Communications and Networking，2008.

[4] A. Banerjee，K. Venkatasubramanian and S. K. S. Gupta，Challenges of Implementing Cyber-Physical Security Solutions in Body Area Networks[C]，In Proc. of BodyNets09，Los Angeles，CA，April 1-3，2009.

本章参考文献

[1] IEEE 802.15 WPAN TG6 Body Area Networks[OL]，http://ieee802.org/15/pub/TG6.html.

[2] 刘艳丽. 基于人体环境的无线体域网网络结构研究[D]. 上海交通大学硕士学位论文，2008.

[3] 宫继兵，王睿，崔莉. 体域网 BSN 的研究进展及面临的挑战[J]. 计算机研究与发展. 47(5)：737-753，2010.

[4] M. Naor，G. Segev，Public-Key Cryptosystems Resilient to Key Leakage[C]，In Proc. of Advances in Cryptology (CRYPTO09)，18-35，2009.

[5] HIPPA 条例，http://www.hhs.gov/ocr/hipaa/.

[6] Dan Boneh，http://crypto.stanford.edu/～dabo/.

[7] S. Saleem，S. Ullah，H. S. Yoo，On the Security Issues in Wireless Body Area Networks[J]，International Journal of Digital Content Technology and its Applications，3(3)：178-184，2009.

[8] Y. Law，J. Doumen，and P. Hartel. Survey and Benchmark of Block Ciphers for Wireless Sensor Networks [TR]. Technical Report TR-CTIT-04-07，Centre for Telematics and Information Technology，University of Twente，The Netherlands，2004.

[9] C. Y. Poon，Y. T. Zhang，and S. D. Bao，A Novel Biometrics Method to Secure Wireless Body Area Sensor Networks for Telemedicine and M-health[J]，IEEE Communications Magazine，44(4)：73-81，2006.

[10] S. D. Bao，C. C. Y. Poon，Y. T. Zhang，and L. F. Shen，Using the Timing Information Of Heart Beats As An Entity Identifier To Secure Body Sensor Network [J]，IEEE Transactions on Information Technology in Biomedicine，12(6)：772-779，2008.

[11] K. Venkatasubramanian，A. Banerjee，and S. K. S. Gupta. Plethysmogram-based Secure Inter-Sensor Communication in Body Area Networks[C]. In Proc. of Military Communications

Conference (MILCOM08), IEEE, 1-7, 2008.
[12] K. Venkatasubramanian, A. Banerjee, and S. Gupta, Ekg-based Key Agreement In Body Sensor Networks[C], In Proc. of INFOCOM Workshop, 1-6, 2008.
[13] M. Raza Kanjee, K. Divi, and H. Liu, A Physiological Authentication Scheme in Secure Healthcare Sensor Networks[C], In Proc. of IEEE SECON10, 1-3, 2010.

第 13 章　M2M 安全

首先 M2M 利用远程无线通信接入到互联网是非常方便快捷的、容易部署、价格不贵、还支持移动性。设想一下，将家里的热水器装上支持 GPRS 网络的 SIM 卡，就可以在下班的路上给它发短信控制其开关和温度了。传感器结点(或者传感器网络的网关)也可以通过 SIM 卡将监控或者感知的数据发给你，如室温，湿度等。设想一下，如果 SIM 采取的是 TD-SCDMA 甚至是数据更快的 4G LTE 网络，(第 6 章曾经提到，TD-LTE 下行速率 100Mbps，上行速率 50Mbps)，那么甚至可以看到视频摄像头监控的内容，这无疑相当于你具备了无数个“千里眼”遍布世界各地，你便可以“感知地球”了。甚至更加大胆的设想，如果你的小车装有 M2M 的 SIM 卡，你给汽车发条短信，它就可以自动开过来接你，因为你的车有卫星导航，且能够感知路前方的障碍物，小车综合它感知的数据便可以实现自动驾驶。

其次，M2M 具有巨大的商机。据 2012 年 2 月的媒体报道，市场研究机构 Machina Research 公布的最新数据显示，排名 M2M(机器与机器)通信领域前 20 位的移动运营商有望于 2020 年从 M2M 应用程序中获得 250 亿欧元的收益。Machina Research 公司给出的前三名分别为沃达丰、德国电信和 AT&T，这三家运营商将占排名前 20 位的运营商总收入的 35%。VerizonWireless、Telefonica 和中国移动居第二集团，分别以微弱差距据第四、第五和第六位，第三集团包括中国联通、Orange、Sprint 和 Telenor 公司(第六至十位)。

同时，M2M 在学术界也引起更多研究者的注意。就在 2011 年，很多学术期刊组织了关于 M2M 的专题讨论(专刊)，关于这一主题的大型学术会议下的专题研讨会(workshop)也非常多。

另外值得注意的是，欧洲电信标准化协会(ETSI)2012 年初发布了其第一版 M2M 标准——ETSI M2M 标准 Release 1。该标准允许多种 M2M 技术之间通过一个可管理平台进行整合。标准对 M2M 设备、接口网关、应用、接入技术及 M2M 业务能力层进行了定义。同时，提供了安全、流量管理、设备发现及生命周期管理特性。该标准的发布被视为是 M2M 产业的一个里程碑，预示着产业发展迈向规范化和规模化。

13.1　M2M 概述

13.1.1　M2M 的概念、架构与应用

M2M(Machine to Machine)即“机器对机器”的缩写，扩展的概念包括“Machine to Mobile”——机器对移动设备、“Machine to Man”——“机器对人”等，旨在实现人、设备、

系统间无缝连接。可认为是机器与机器、人与机器、移动网络与机器之间的连接与通信，涵盖了所有实现人、机器、系统之间建立通信连接的技术和手段。本节主要讨论狭义的概念，即通过无线移动通信网络进行广域网连接的设备间的通信。

M2M 强调机器与机器的互联，连接方式是多种多样的。但是，移动通信网络由于其网络的特殊性，终端侧不需要人工布线、可以提供移动性支撑，有利于节约成本，并可以满足在危险环境下的通信需求，这使得以移动通信网络作为承载的 M2M 服务得到了业界的广泛关注。正是由于上述原因，M2M 通常是由电信部门（移动运营商）主导和助推的。图 13.1 给出了一个 M2M 系统结构框图。

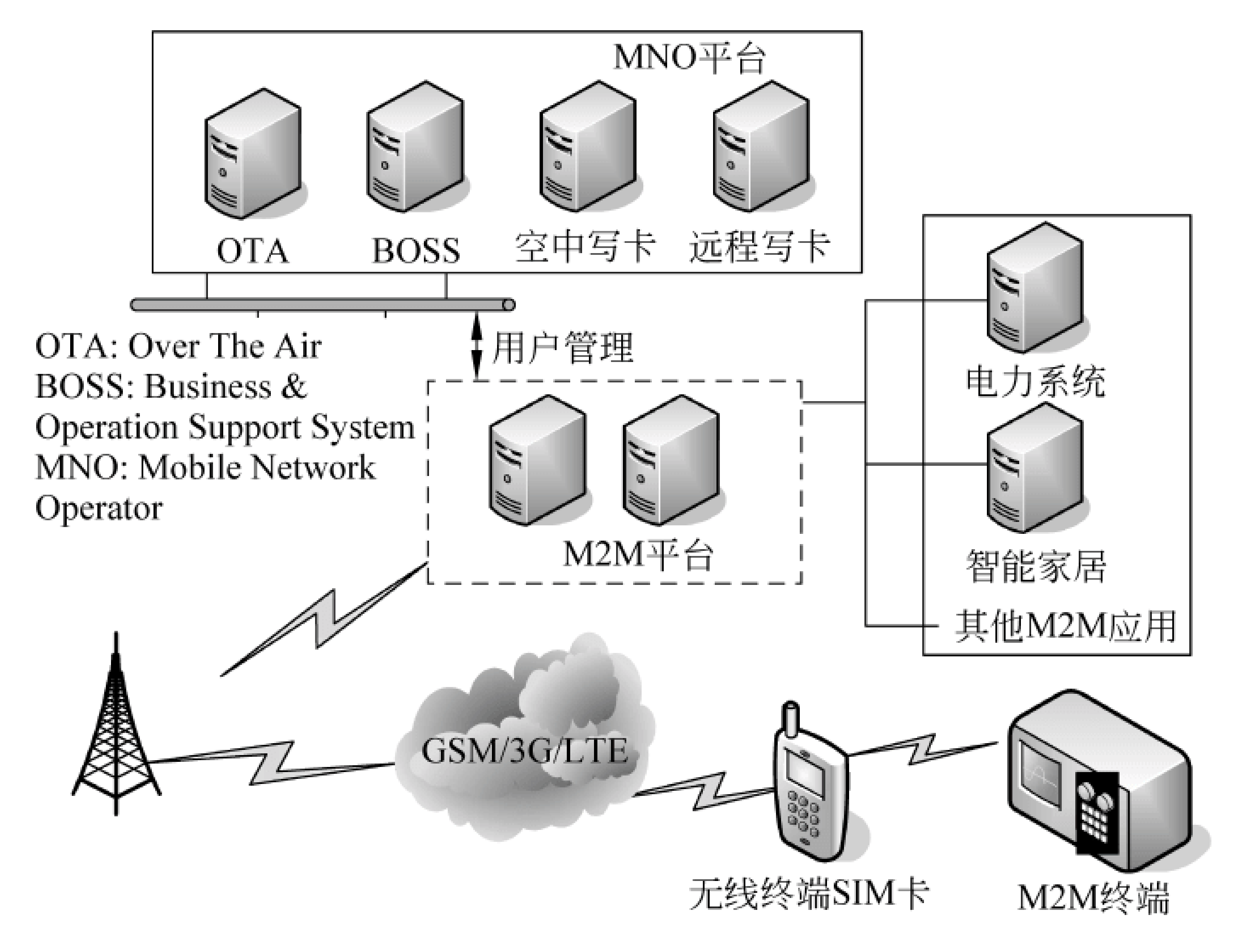


图 13.1　M2M 系统结构框图

早在第一章我们就指出，M2M 可以和 EPCglobal 网络融合，即 RFID 读写器通过 M2M 连接到 Internet，然后可访问 EPGglobal 定义的 ONS（Object Name Service）、EPCIS（EPC Information Service）等服务。M2M 和 6LoWPAN 技术之间的区别是 6LoWPAN 提供了直接的 Internet 寻址能力，而 M2M 是通过在 M2M 服务器端的网关功能进行寻址，这种寻址类似于一种基于广域无线通信网的网络地址转换（NAT，Natural Address Translatoin），因为 M2M 可不需要配置 IP 地址，只需要配置内部标识，即 MCIM（M2M Communications Identity Module）。

欧洲电信标准委员会 ETSI（European Telecommunications Standards Institute）研究的 M2M 相关标准有十多个，其中一个提出了 M2M 的功能体系架构，包括定义新的功能实体，与 ETSI 其他分支或其他标准化组织标准间的标准访问点和概要级的呼叫流程。M2M 的体系架构见图 13.2，可以看出，广义地讲 M2M 技术涉及了通信网络中从终端到网络再到应用的各个层面，M2M 的承载网络包括了 3GPP，TISPAN（Telecommunications and Internet converged Services and Protocols for Advanced Networking，电信和互联网融合业务及高级网络，TISPAN 是 ETSI 从事下一代网络 NGN 研究的标准化组织，由欧洲运营商和制造商主导），以及 IETF 定义的多种类型的通

信网络。

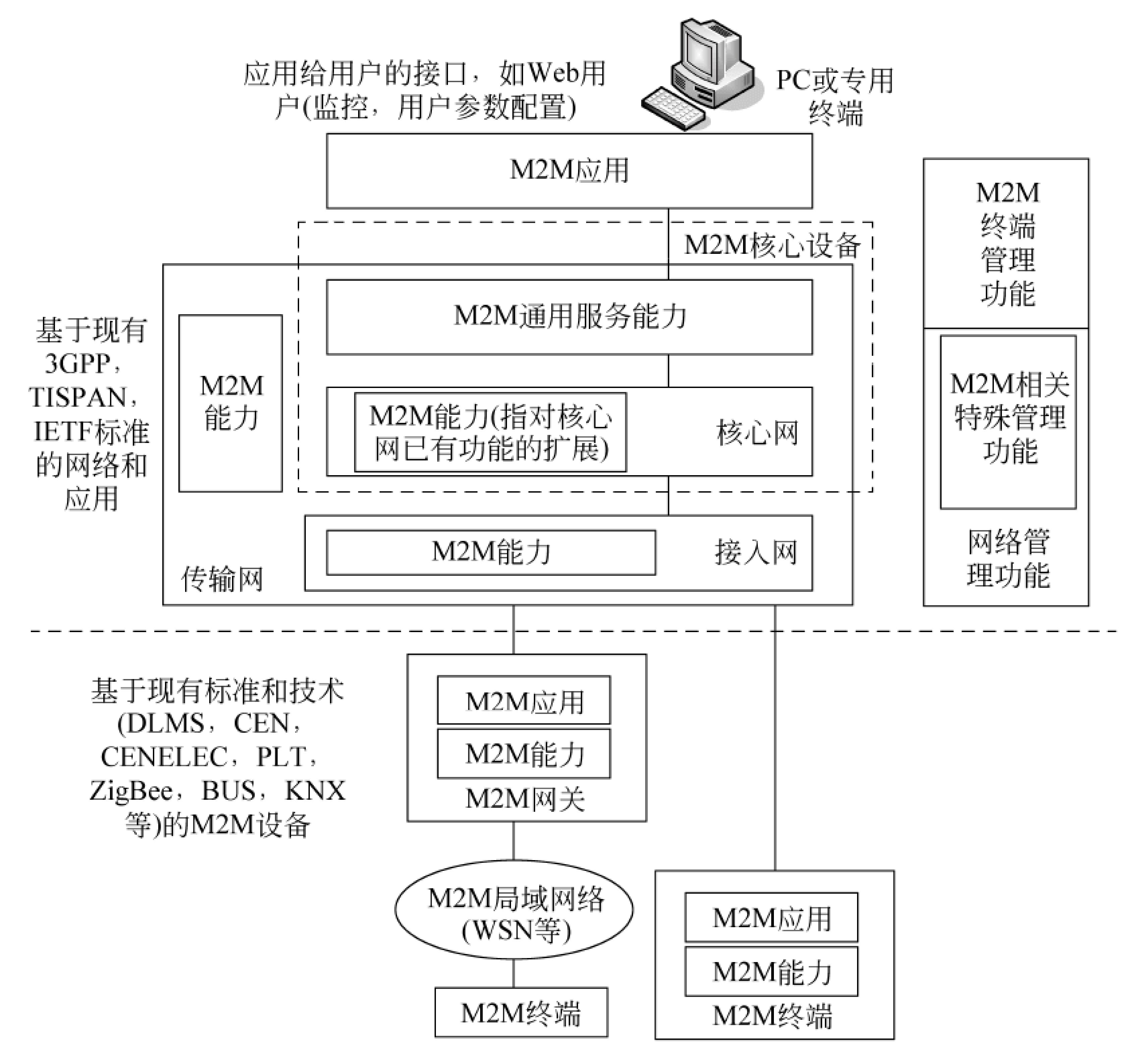


图 13.2　ETSI M2M 功能架构中的针对机器通信的功能强化

M2M 技术在欧美、韩国和日本实现了商用，美国的 Sprint 公司通过 ODI 计划认证了 160 个厂商的 M2M 终端设备，这些设备被广泛应用到智能抄表、无线 POS 机、车载管理等多个行业领域。到 2009 年 9 月，日本国际电信电话株式会社 KDDI 已经售出 200 万台 M2M 通信模块，主要针对车辆管理行业和儿童定位行业。

关于我国在 M2M 技术方面，中国移动已经开通了 M2M 业务，推出了"车务通"、"电梯卫士"、"消防监控系统"、"爱贝通"、"关爱通"5 项 M2M 应用。2007 年起，M2M 杂志每年定期发布 M2M 领域内的全球 100 强名单(包括 M2M 产业链上下游企业)的成员，获得全球 M2M 业的广泛认可。全球 M2M 模块主要供应商有 Cinterion(从西门子公司分出)、Wavecom 和 Telit。

我国 M2M 模块的主要供应商是 SIMCom(芯讯通)公司。据媒体报道 2010 年 SIMCom 公司以 18%的份额稳居全球无线模块市场第二的位置，成为全球 M2M100 中的唯一国内模块厂家。2010 年 6 月，该公司推出新的 TD-SCDMA 模块 SIM4200。这是一款专为中国移动网络设计的工业级产品，该产品同时支持中国移动的 WMMP 协议。SIM4200 是一款双模 TD-SCDMA/HSDPA/EDGE/GSM 模块，其中 TD-SCDMA/HSDPA 工作频段为 1880～1920MHz 及 2010～2025MHz。其数据传输能力可达到下行速度 2.8Mbps，上行速度 384Kbps。SIM4200 采用板对板的封装形式，支持工业应用程

序常用的 USB2.0 和串行端口(RS232)接口,支持语音。另外,该产品还支持 Java 二次开发应用和嵌入了中国移动 WMMP(Wireless Machine To Machine Protocol)协议。

WMMP 协议是中国移动为实现终端与 M2M 平台数据通信过程而设计,建立在 UDP 协议之上。嵌入 WMMP 协议到通信模组内,可以提供标准化的软硬件接口以及二次应用开发环境,实现物联网终端的标准化和快速部署。

M2M 的应用遍布各个领域,包括交通领域(物流管理、定位导航、智能交通等)、电力领域(远程抄表和负载监控)、农业领域(大棚监控、动物溯源)、城市管理(电梯监控、路灯控制)、安全领域(城市和企业安防)、环保(污染监控、水土检测)、企业(生产监控和设备管理)和家居(老人和小孩看护、智能安防)等。M2M 具有巨大的市场潜力和应用前景,是物联网的主要应用支撑技术。

特别是国内最近提出的"车联网"概念,可以应用 M2M 技术,直接在车辆上安全 M2M 设备,通过 GPRS(TD-SCDMA、LTE)接入 Internet。这个接入规模是巨大的,如果再和 GPS(或者北斗系统)功能连用,能导致很多新的应用。注意,在无线网络的学术界,还有一种车辆间的无线网络叫做车载自组织网络,即 VANET(Vehicular Ad-hoc Network),这种网络依赖独立的建网、使用的不同的物理层协议,需要路边单位(Road Side Unit)的支持,已经研究了多年,在北美讨论较多,但因各国国情不同,一直没有广泛应用。

国际上各大标准化组织中 M2M 相关研究和标准制定工作也在不断推进。几大主要标准化组织按照各自的工作职能范围,从不同角度开展了针对性研究。

(1) 欧洲电信标准化委员会(ETSI)是欧盟建立的非盈利性的电信标准化组织,成立的 M2M 技术委员会(M2M TC)。从典型物联网业务用例,例如智能医疗、电子商务、自动化城市、智能抄表和智能电网的相关研究入手,完成对物联网业务需求的分析、支持物联网业务的概要层体系结构设计以及相关数据模型、接口和过程的定义,M2M 解决方案间的互操作性等。

(2) 3GPP/3GPP2 以移动通信技术为工作核心,重点研究 3G,LTE/CDMA 网络针对物联网业务提供而需要实施的网络优化相关技术,研究涉及业务需求、核心网和无线网优化、安全等领域。

(3) TISPAN 于 2004 年开始 NGN 全球标准的研究,目前 TISPAN 分为 8 个研究组和一个系统组,分别为 WG1 业务、WG2 体系架构、WG3 协议、WG4 编号和路由、WG5 家庭网络、WG6 测试、WG7 安全、WG8 网络管理,系统组负责各组共同相关的问题或协调事宜。

(4) 中国通信标准化协会 CCSA 早在 2009 年完成了 M2M 的业务研究报告,与 M2M 相关的其他研究工作已经展开。我国 M2M 相关的标准化工作由 CCSA 的移动通信工作委员会(TC5)和泛在网技术工作委员会(TC10)负责。主要工作内容如下:

① TC5 WG7 完成了移动 M2M 业务研究报告,描述了 M2M 的典型应用、分析了 M2M 的商业模式、业务特征以及流量模型,给出了 M2M 业务标准化的建议。

② TC5 WG9 于 2010 年立项的支持 M2M 通信的移动网络技术研究,任务是跟踪 3GPP 的研究进展,结合国内需求,研究 M2M 通信对 RAN 和核心网络的影响及其优化

方案等。

③ TC10 WG2 M2M 业务总体技术要求，定义 M2M 业务概念、描述 M2M 场景和业务需求、系统架构、接口以及计费认证等要求。

④ TC10 WG2 M2M 通信应用协议技术要求，规定 M2M 通信系统中端到端的协议技术要求。

可以看出，关于我国 M2M 的安全标准化工作还需要完善。

13.1.2 M2M 应用实例

在现有阶段，M2M 应用通常依赖于 GSM 模块，将来会鼓励更多地使用 TD-SCDMA 模块。GSM 模块根据提供的传输速率可分为 GPRS 模块、EDGE 模块和纯短信模块。短信模块只支持语音和短信服务。GPRS 的传输速率从 56Kbps 到 114Kbps 不等，理论速率为 171Kbps。比 GSM 9.6Kbps 的传输速率高。GPRS 技术还具有在任何时间、任何地点都能实现连线、永远在线、按流量计费等特点。EDGE 技术进一步提升数据传输的速率到 384～473Kbps。目前国内的 GSM 网络普遍具有 GPRS 通信功能，移动和联通的网络都支持 GPRS，而 EDGE 在部分省市实现了网络覆盖。

GPRS 模块是具有 GPRS 数据传输功能的 GSM 模块。其实就是一个精简版的手机，如果加上屏幕和键盘就是一个完整的手机。单片机（如 ARM 机）可通过 RS-232 串口与 GPRS 模块相连，通过 AT 指令控制 GPRS 模块实现通信功能。市场上比较流行的模块包括西门子的 TC35i、MC39i、MC55，Wavecom 的 Q24/Q26 系列，大唐的 B200、B255，SIMCom 公司的 SIM300 等。对于模块的应用，主要掌握软件和硬件两个方面。软件主要是了解模块的 AT 指令集，包括标准的 AT 指令和各厂家自己扩展的 AT 指令，例如 TCP/IP 指令、STK 指令等。硬件主要是了解该模块的硬件接口，如串口、SIM 接口等，电气参数如何，如电压、电流要求等。

下面给出一个 M2M 应用的实例[4]。某省无人值守的气象信息采集系统。各气象采集点通过 GPRS 模块与中心站主机保持实时连接，监测设备获得的气象信息通过 GPRS 模块传给气象中心主机，中心主机对信息处理，形成气象预报信息。GPRS 模块提供了广域无线 IP 连接，设备安装方便、维护简单、缩短了建设周期、降低了建设成本。采集点的覆盖范围广，接入地点无限制，扩容无限制（因为 GPRS 已经覆盖该省绝大部分地区）。数据传输速率高，通信费用低（GPRS 按流量计费，每月通信费约 200 元之内）。具有良好的实时响应和处理能力。

系统的构成包括：

（1）气象信息采集点，采用 G200 GPRS 模块，通过 RS-232/RS-485/TTL 接口与气象设备采集点连接，接入中国移动提供的 GPRS 网。数据采集点使用了专用的 STK 卡，该卡只能用于与中心站得数据通信。

（2）气象中心站。维护接入的每个终端的 IP 地址和 ID 号，确保对采集点的寻址。系统还支持各种行业应用，如实现信用卡实时认证、POS 的远程控制维护、远程业务点接入等，如图 13.3 所示。

G200 GPRS 模块具有如下特点：

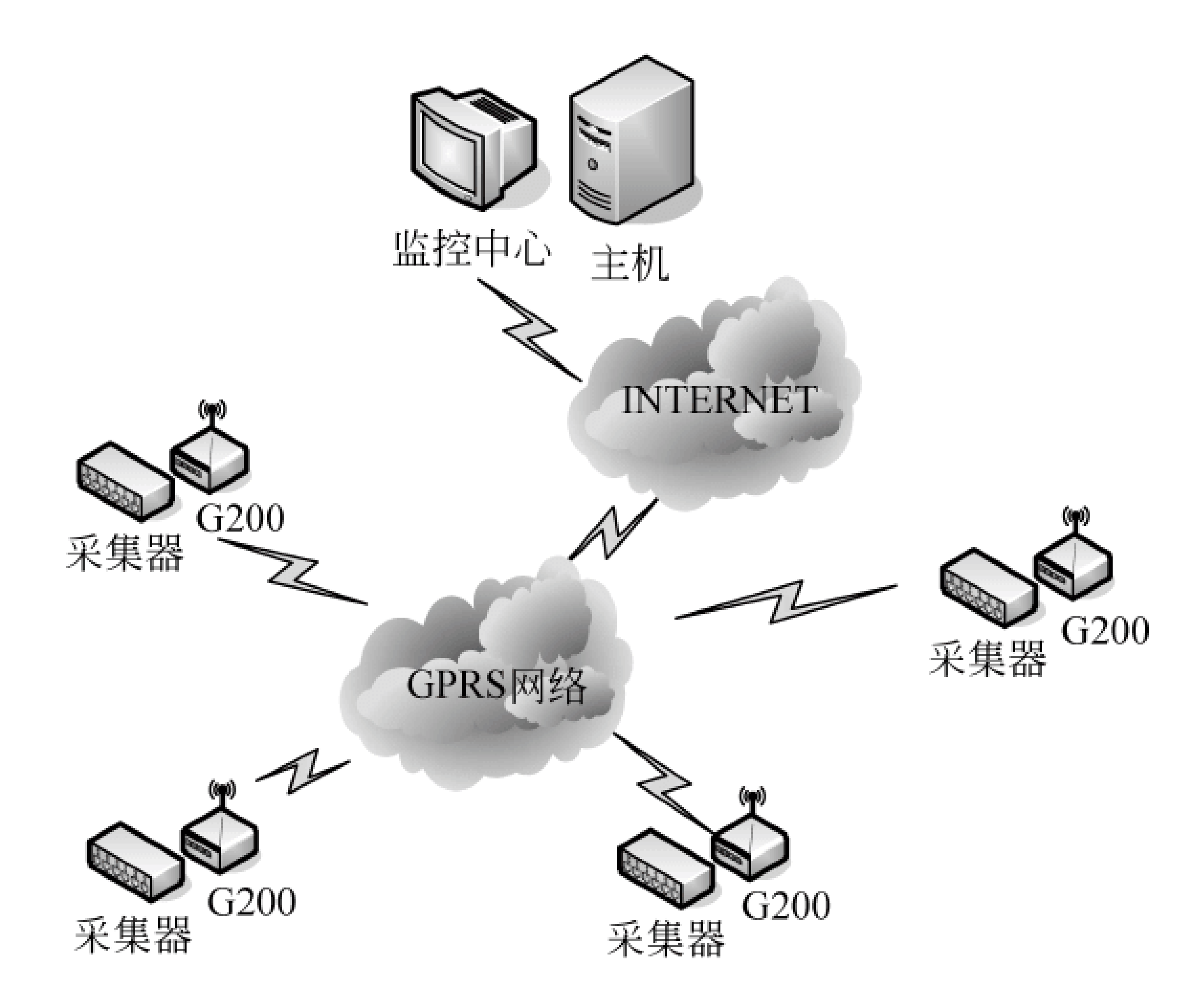


图 13.3　基于 GPRS 模块的无人值守的气象信息采集系统

① 支持 900/1800 双频；

② 接口支持 RS-232，RS-485，TTL；

③ 理论传输速率为 171Kbps，实际传输速率为 40Kbps；

④ 支持 Windows 98/2000/XP/Linux 操作系统；

⑤ 内嵌 TCP/IP 协议，可完成透明传输通道；实时监测网络连接情况，掉线自动重传；

⑥ 提供主/副 IP 及动态域名解析；

⑦ 用户可设定心跳报告时间间隔；

⑧ 用户可设定数据通信帧长度等。

上述实例其实可代表大部分的 M2M 应用模式，它采用的是直接将传感器结点通过 M2M 方式传递到 GPRS 网络，经过 Internet 连接到控制中心。如果在终端具有多个传感器结点，如构成了传感器结点的簇结构，可以先通过 ZigBee 将这些结点的数据传送到簇头结点，也称为 Sink 结点或网关结点，由网关结点通过 M2M 方式发送到 GPRS 网络，经 Internet 到控制中心。

如果进一步扩展，终端结点可加上 GPS(或北斗系统)模块、RFID 模块、甚至是传感器和制动器模块的组合，便可以构成带有定位、标识、感知、行动四种功能的应用。其网络接入的实现方法是类似的。

其实，这一实例系统中需要保障系统安全和稳定，包括信道加密、信源加密、登录保护、访问防护、接入防护、防火墙等。通常此类系统的安全防护可以包括：

(1) 利用 SIM 卡的唯一性，对用户 SIM 卡手机号码进行鉴别授权。在网络中对 SIM 卡号和 APN(Access Point Name)进行绑定，划定用户可以接入某系统的范围，只有属于指定行业的 SIM 卡手机号才能访问专用 APN。

这里 APN 是指“接入点名称”，是手机上网时必须配置的一个参数，它决定了手机通过哪种接入方式来访问网络，通常用来标识 GPRS 的业务种类，目前国内分为两大类：通

过 GPRS 访问 WAP 业务 CMWAP/UNIWAP/3GWAP;以及(用上网卡)访问 Internet 的服务 CMNET/UNINET/3GNET。两种服务的资费不同。

(2) 对于特定用户,可通过数据中心分配特定的用户 ID 和密码,没有分配 ID 和密码的用户将无法登录系统。

(3) 数据加密。通过 VPN 对整个数据传送过程进行加密。

(4) 网络接入安全鉴权。采用防火墙软件,设置网络鉴权和安全防范功能。

13.2 M2M 安全

13.2.1 M2M 的安全威胁与对策

1. 本地安全威胁

M2M 通信终端很少是有人直接参与管理的,因而存在许多针对 M2M 终端设备和签约信息的攻击[5,6]。

(1) 盗用 M2M 设备或签约信息。

M2M 设备在一般情况下是无专人看管的,这就可能导致攻击者、盗取 USIM (Universal Subscriber Identity Module)或 UICC(Universal Integrated Circuit Card)甚至 M2M 设备。从而窃取或篡改 M2M 设备中的用户签约信息。

防御这一威胁的方法是采取机卡一体方案,承载 USIM 应用的 UICC 不可被移除或移除后将永久失效;在机卡分离的方案中,移除 UICC 将导致 M2M 设备不可用;对于将 MCIM(M2M Communications Identity Module)直接以软件形式绑定到 M2M 设备中的方案,M2M 设备则需要提供一种特殊的可信环境来安全存储和执行 MCIM,例如可信计算技术中的远程证明机制,同时能够防止攻击者通过逻辑或物理方式进行攻击。

(2) 破坏 M2M 设备或签约信息。

攻击者可能会采用物理或逻辑方法改变 TRE(TRusted execution Environment)的功能、TRE 与 M2M 设备间的控制信息或已获取的 MCIM 中的信息,造成用户无法接入网络或丢失个人数据等。或者直接破坏 M2M 设备或 UICC,造成签约信息或 M2M 设备不可用。攻击者还可以将 MCIM 的承载实体暴露在有害电磁环境中,导致其受到破坏从而造成签约信息不可用。此外,攻击者还可以通过向 M2M 设备中添加恶意信息导致签约信息不可用。

可采取的对策是 M2M 设备应具备较强的抗辐射、耐高低温能力,为 M2M 设备中的功能实体提供可靠的执行环境。

2. 无线链路的安全威胁

M2M 终端设备与服务网之间的无线接口可能面临以下威胁。

(1) 非授权访问数据。

攻击者可以窃听无线链路上的用户数据、信令数据和控制数据,进行流量分析,从而获取 M2M 用户密钥或控制数据等机密信息,非法访问 M2M 设备上的数据。采取的对

策是在终端设备与服务网之间使用双向认证机制以及采取相应的数据加密算法。

（2）对完整性的威胁。

攻击者可以修改、插入、重放或者删除无线链路上传输的合法 M2M 用户数据或信令数据，对 M2M 用户的交易信息造成破坏。这可以通过完整性保护方法如使用具有完整性密钥的 Hash 算法来保护。

（3）拒绝服务攻击。

攻击者通过在物理层或协议层干扰用户数据、信令数据或控制数据在无线链路上的正确传输，实现无线链路上的拒绝服务攻击。拒绝服务攻击是攻击无线链路的常见攻击，目前没有特别好的方法，有些人提出基于物理层（如数据链路层）的解决方法，但适用性可能不高。通常可采取的对策是采取追踪机制确定攻击者的位置。

3. 对服务网的安全威胁

服务网通常是由通信网的支撑网组成。通常通信网可分为业务网、传输网和支撑网，其中支撑网包括第 7 号信令网、数字同步网、智能网和电信管理网。

（1）非授权访问数据和服务。

攻击者可以进入服务网窃听用户数据、信令数据和控制数据，未经授权访问存储在系统网络单元内的数据，进行流量分析。

攻击者冒充合法用户使用网络服务，通过改变 MCIM 的接入控制方式来获得非法服务，利用窃取的或过期未注册的身份来注册获取 MCIM，从而获得非授权的签约信息接入服务网获取服务。攻击者可能假冒成归属网以获取假冒另一用户所需的信息。解决上述问题的方式是采用认证机制。

（2）终端病毒或恶意软件。

攻击者可以通过恶意软件、木马程序或其他手段获取 M2M 上的应用软件、签约信息以及 MCIM，然后在其他 M2M 设备上复制还原，从而冒用 M2M 用户的身份；还可以通过病毒或恶意软件更改、插入、删除用户的通信数据。采取的对策是 M2M 设备应能够定期更新防病毒软件，若采取远程软件更新需经过签名，M2M 设备应能验证远程更新的软件的合法性。

13.2.2　M2M 的安全标准和研究进展简介

在国际标准的安全方面，3GPP 给出了一个报告。M2M 在 3GPP 内对应的名称为机器类型通信（MTC，Machine-Type Communication），其 SA3 工作组负责安全相关的研究。该工作组发布的最新的版本是 2010 年 6 月发布的文档 SP-48 3GPP：3GPP TR 33.812 V9.2.0[7]，即《M2M 设备签约信息的远程提供和变更中安全方面的可行性研究》，英文名为 Feasibility study on the security aspects of remote provisioning and change of subscription for Machine to Machine（M2M）equipment，该文档讨论了 M2M 应用在 UICC 中存储时，M2M 设备的远程签约管理，包括远程签约的可信任模式、安全要求及其相应的解决方案等。

ETSI 于 2011 年 1 月给出了一个标准草案，Draft ETSI TR 103 167 V0.2.1，即 M2M：Threat Analysis and Counter-Measures to M2M Service Layer（M2M 业务层威胁

分析及对策)。在最近一次研讨会(2012 年初)上指出 M2M 业务层安全架构中将在 M2M 设备和网关上支持双向认证,完整性保护和保密性[8,9]。

M2M 的安全在学术界的研究还刚起步,下面加以简要介绍。

文献[10]指出 M2M 通信中的挑战是能量效率(Green),可靠性(Reliability)和安全(Security),缩写为 GRS。没有 GRS,M2M 通信将不可能广泛采用。

文献[11]综述了 M2M 中的安全和可信问题。从应用场景用例(Use Case)入手,结合 3GPP 安全工作组 SA3 对安全威胁的分析,安全威胁包括物理攻击(physical attacks),包括将有效的卡插入到一个特制的设备,用虚假的或者修改的软件启动初始化过程,环境或侧信道攻击等,对 SIM 卡的攻击(compromise of credentials),如对卡中认证算法的攻击,对卡的物理入侵,侧信道攻击,以及卡的克隆。配置攻击(configuration attacks),如虚假的软件更新和配置改变,对访问控制表的错误配置等。协议攻击(protocol attacks),包括中间人攻击,拒绝服务攻击,对 OTA(Over-The-Air)管理的攻击。对核心网的攻击(attacks on the core network),主要指对 MNO(Mobile Network Operator)的威胁,包括设备的伪造,在伪造设备间建立流量隧道(traffic tunneling),在路由器或者网关上错误地配置防火墙,对核心网拒绝服务攻击,非授权地改变设备的认证的物理位置,使用虚假设备攻击网络等。用户数据以及身份标识的隐私攻击(user data and identity privacy attacks),包括偷听用户和设备的通过接入网发送的数据,伪装成另一个用户或者签约设备,透露用户的网络 ID 或者机密数据给非授权第三方。

文献[12]指出 M2M 中,以车辆安全模块(Vehicle Security Module)为例,该模块有两个功能,通过手机通信网络进行车辆解锁和发动引擎。文章作者证明了攻击者可以非常容易地对该模块发布这两个命令,绕过认证和授权环节。模块在发布这两个命令时,需要访问 CAN(Controller-Area Network),即控制域网络,这是一个车辆内部组件之间的通信网络。发送给 CAN 总线的命令需要被认证,即车辆安全模块要求认证的发送者才能发送消息给 CAN 总线。若攻击者可以攻陷一个模块,则可以用该认证模块的名义发送消息。从而攻击者可以控制攻陷模块以外的组件,如灯、刹车、油门等。可见,M2M 在提供给用户远程控制的方便之余,也对安全提出了更高的要求。

最后,关注一下商业方面的进展作为实际的应用需求参考。国内的握奇公司提出的 M2M 平台功能架构中包括一个安全访问控制模块,这个模块的功能主要针对码号资源管理,SIM 个人化、密钥管理和鉴权访问控制。对于运营商来说,当前码号资源十分紧张,如何在物联网时代,解决码号资源问题就成为当务之急。为了解决码号资源问题和安全管理,该公司推出了空中发卡解决方案和机卡认证解决方案。空中发卡解决方案利用了虚拟号码池和码号资源回收机制实现码号资源充分利用,避免了号码资源闲置浪费的局面。安全认证解决方案利用了终端的 IMEI 号和 SIM 模块的 IMSI 进行机卡绑定,同时通过 M2M 鉴权认证平台进行机卡互锁管理,平台定期下发更新密钥,可防止盗卡、盗机现象,确保码号资源的安全性。

总之,M2M 的安全研究才刚刚开始,还有很多安全问题等待着去发现和解决。

研究与思考

[1] 本章参考文献[4]给出了物联网通信模块的芯片级设计指导,其中有很多 M2M 应用的实际例子。尝试利用该书开发一个实用的模块作为练习。
[2] 使用 SIMCom 公司 SIM4200 模块开发一个车联网的应用。思考其中可能会遇到的网络安全问题。
[3] M2M 的安全性往往和移动通信网络的安全性有关联,M2M 模块可视为精简版的智能手机,那么在安全性方面是否可以精简呢。

进一步阅读建议

IEEE 杂志给出了关于某个新的论题的概述,由于其受众面涵盖了学术界和工业界,因而文章相对汇刊而言比较容易读懂,而且论文的发表周期相对较短,对新论题的介绍比较及时。

[1] Geng Wu; S. Talwar, K. Johnsson, N. Himayat, K. D. Johnson, M2M: From mobile to embedded internet[J], IEEE Communications Magazine, vol. 49, no. 4, 36-43, April 2011.
[2] L. Foschini, T. Taleb, A. Corradi, D. Bottazzi, M2M-based metropolitan platform for IMS-enabled road traffic management in IoT[J], IEEE Communications Magazine, vol. 49, no. 11, 50-57, November 2011.
[3] Yan Zhang, Rong Yu, Shengli Xie, Wenqing Yao, Yang Xiao, M. Guizani, Home M2M networks: Architectures, standards, and QoS improvement[J], IEEE Communications Magazine, vol. 49, no. 4, 44-52, April 2011.
[4] Chang Kim, A. Soong, M Tseng, Zhixian Xiang, Global Wireless Machine-to-Machine Standardization[J], IEEE Internet Computing, vol. 15, no. 2, 64-69, March-April 2011.

本章参考文献

[1] 分析称移动运营商或在 M2M 市场创收 250 亿欧元[OL],http://tech.qq.com/a/20120224/000309.htm.
[2] M2M magazine[OL], http://www.machinetomachinemagazine.com.
[3] M2M 技术在国内外标准进展现状[J]. 金卡工程. 2011 年 11 月.
[4] 夏华. 无线通信模块设计与物联网应用开发[M]. 北京:电子工业出版社,2011.
[5] 晁世伟,杨元,李静毅. 物联网 M2M 的安全分析及策略[J]. 计算机科学,2011.10,38(10A):7-9.
[6] 焦文娟,齐旻鹏,朱红雪. M2M 的安全研究[J]. 电信技术,2009.6,76-78.
[7] 3GPP TR 33.812, Feasibility study on the security aspects of remote provisioning and change of subscription for Machine to Machine (M2M) equipment[S]. http://www.3gpp.org/ftp/Specs/html-info/33812.htm.
[8] ETSI TC M2M Security Architecture, ETSI workshop[OL], Jan. 2012, http://workshop.etsi.

org/2012/201201_SECURITYWORKSHOP/REPORT. pdf.
[9] Alper Yegin，ETSI M2M TC WG4，ETSI M2M Security Architecture[OL]，http://workshop. etsi. org/2012/201201 _ SECURITYWORKSHOP/7 _ M2MandSMARTGRIDS/ETSITCM2M _ Security_Architecture_YEGIN%20. pdf.
[10] Rongxing Lu，Xu Li，Xiaohui Liang，Xuemin Shen，Xiaodong Lin，GRS：The green，reliability，and security of emerging machine to machine communications[J]，IEEE Communications Magazine，vol. 49，no. 4，28-35，April 2011.
[11] C Inhyok，Y. Shah，A. U. Schmidt，A. Leicher，M. V. Meyerstein，Trust in M2M communication[J]，IEEE Vehicular Technology Magazine，vol. 4，no. 3，69-75，Sept. 2009.
[12] D. A. Bailey，Moving 2 Mishap：M2M's Impact on Privacy and Safety[J]，IEEE Security & Privacy，vol. 10，no. 1，84-87，Jan. -Feb. 2012.

全书参考文献[①]

［1］ 周洪波. 物联网：技术、应用、标准和商业模式［M］. 第 2 版. 北京:电子工业出版社，2011.
［2］ 杨刚，沈沛意，郑春红等. 物联网理论与技术［M］. 北京:科学出版社，2010.
［3］ 刘云浩. 物联网导论［M］. 北京:科学出版社，2011.
［4］ 王汝传，孙力娟. 物联网技术导论［M］. 北京：清华大学出版社，2011.
［5］ J. P. Vasseur，A. Dunkels，田辉，徐贵宝，马军峰等译. 基于 IP 的物联网架构、技术与应用［M］. 北京：人民邮电出版社，2011.
［6］ 刘海涛. 物联网：技术应用［M］. 北京：机械工业出版社，2011.
［7］ 任伟. 无线网络安全［M］. 北京：电子工业出版社，2011.
［8］ 任伟. 现代密码学［M］. 北京：北京邮电大学出版社，2011.
［9］ 任伟译. J. Katz. Y. Lindell. 现代密码学——原理与协议［M］. 北京：国防工业出版社，2010.
［10］ 任伟译. J. Katz，数字签名［M］. 北京：国防工业出版社，2012.

① 这里列出的主要是在全书多处章节用到的参考文献，另外每章参考文献在每章已经列出，方便只阅读部分章节的读者查阅。